

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 986 412**

51 Int. Cl.:

C12N 15/10 (2006.01)

C12Q 1/6886 (2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **27.10.2017** **PCT/US2017/058789**

87 Fecha y número de publicación internacional: **03.05.2018** **WO18081580**

96 Fecha de presentación y número de la solicitud europea: **27.10.2017** **E 17863995 (1)**

97 Fecha y número de publicación de la concesión europea: **10.07.2024** **EP 3532637**

54 Título: **Método de aislamiento de ácidos nucleicos eucariotas a partir de muestras de heces**

30 Prioridad:

27.10.2016 US 201662413708 P

22.06.2017 US 201762523511 P

17.08.2017 US 201762547046 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
11.11.2024

73 Titular/es:

GENEOSCOPY, INC. (100.0%)

2220 Welsch Industrial Ct.

St. Louis, MO 63146, US

72 Inventor/es:

BARNELL, ERICA;

BARNELL, ANDREW;

KANG, YIMING y

WURTZLER, ELIZABETH

74 Agente/Representante:

SÁEZ MAESO, Ana

Observaciones:

Véase nota informativa (Remarks, Remarques o Bemerkungen) en el folleto original publicado por la Oficina Europea de Patentes

ES 2 986 412 T3

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

DESCRIPCIÓN

Método de aislamiento de ácidos nucleicos eucariotas a partir de muestras de heces.

Campo de la invención

La presente invención se refiere a la extracción de ácidos nucleicos eucariotas a partir de muestras de heces y al uso de los ácidos nucleicos para el diagnóstico y tratamiento de enfermedades intestinales.

Antecedentes

Los trastornos gastrointestinales, por ejemplo el cáncer gastrointestinal y otras enfermedades digestivas tal como la enfermedad inflamatoria intestinal, el síndrome del intestino irritable y la enfermedad de Crohn están muy extendidos. Se calcula que en Estados Unidos los trastornos gastrointestinales afectan anualmente 60 a 70 millones de personas. En algunos trastornos, la detección y el diagnóstico precoz han permitido reducir las tasas de mortalidad y mejorar la calidad de vida de los pacientes. Sin embargo, los métodos estándar de diagnóstico, tales como la colonoscopia, son invasivos, requieren mucho tiempo y se asocian a costes relativamente elevados. Los trastornos gastrointestinales también pueden afectar a los animales, por ejemplo a los que se tienen como mascotas, como perros y gatos. Los métodos veterinarios de detección de estos trastornos son igualmente invasivos y costosos. Existe una necesidad continua de métodos no invasivos para diagnosticar trastornos gastrointestinales tanto en humanos como en animales. El documento Yu et al. 2016 (CMGH 2(2): 158-174) se refiere a un método de aislamiento de ácido nucleico eucariota a partir de una muestra de heces.

Resumen

La invención se define por el alcance de las reivindicaciones adjuntas. Se proporcionan métodos de acuerdo con las reivindicaciones para aislar ácido nucleico eucariota a partir de una muestra de heces. El método incluye las etapas de mezclar la muestra con un tampón, un tensioactivo y un inhibidor de ribonucleasas para formar una suspensión; separar la suspensión en una porción enriquecida para células eucariotas y una porción enriquecida para células bacterianas y retener la porción enriquecida para células eucariotas; añadir un agente caotrópico y, opcionalmente, un tensioactivo a la porción enriquecida en células eucariotas para formar un lisado; fraccionar el lisado en una capa de restos celulares, una capa que comprende ácidos nucleicos eucariotas y una capa lipídica; y recoger la capa que comprende ácidos nucleicos eucariotas y, opcionalmente, la capa lipídica. La muestra de heces puede ser humana o de un animal no humano. En algunas realizaciones, la muestra de heces de animales no humanos puede ser una muestra obtenida de un perro o un gato. El método puede incluir además la extracción de los ácidos nucleicos eucariotas de la capa recogida que comprende ácidos nucleicos eucariotas. El ácido nucleico puede incluir ADN, ARN, ARN total, ARNm, ARNt, ARNr, ARNnc, ARNsm o ARNsno, o una combinación de cualquiera de ADN, ARN, ARN total, ARNm, ARNt, ARNr, ARNnc, ARNsm o ARNsno.

También se proporcionan métodos para detectar un biomarcador eucariota en una muestra de heces que comprende realizar el método de aislamiento de acuerdo con las reivindicaciones.

El método puede incluir las etapas de analizar el ácido nucleico extraído mediante secuenciación de microarreglos, código de barras molecular, captura de sondas, reacción en cadena de la polimerasa (PCR), ddPCR, RT-PCR, RT-qPCR o secuenciación de ácidos nucleicos. En algunas realizaciones, el biomarcador eucariota se selecciona de entre los biomarcadores listados en la Figura 6 (Panel A) o en la Figura 13 (Panel B). En algunas realizaciones, el biomarcador eucariota puede ser un marcador de células B, un marcador de células T o una inmunoglobulina.

Los métodos para determinar si un sujeto está en riesgo de cáncer colorrectal pueden incluir las etapas de medir el nivel de expresión de dos o más genes biomarcadores de neoplasia colorrectal seleccionados de cualquiera de los genes biomarcadores de neoplasia colorrectal listados en la Figura 6 (Panel A) o Figura 13 (Panel B) en una muestra biológica del sujeto; comparar el nivel de expresión medido de los dos o más genes biomarcadores de neoplasia colorrectal en la muestra con el nivel de expresión medido de los dos o más genes biomarcadores de neoplasia colorrectal en un control, en el que una diferencia en el nivel de expresión medido de los dos o más genes en la muestra biológica en relación con el nivel de expresión medido de los dos o más genes en el control indica que el sujeto está en riesgo de cáncer colorrectal. En algunas realizaciones, los genes biomarcadores de neoplasia colorrectal pueden estar contenidos dentro de los 200 grupos de transcripción expresados diferencialmente mostrados en la Figura 5A y contenidos dentro de las vías comunes asociadas con el cáncer colorrectal mostradas en la Figura 5B. En algunas realizaciones, el gen biomarcador de neoplasia colorrectal puede seleccionarse de entre los biomarcadores listados en la Figura 6 (Panel A) o en la Figura 13 (Panel B).

Los métodos para un plan clínico para un sujeto que tiene o está en riesgo de cáncer colorrectal pueden incluir las etapas de: medir el nivel de expresión de dos o más genes biomarcadores de neoplasia colorrectal seleccionados de cualquiera de los genes biomarcadores de neoplasia colorrectal listados en la Figura 6 (Panel A) o Figura 13 (Panel B) en una muestra biológica del sujeto; comparar el nivel de expresión medido de los dos o más genes biomarcadores de neoplasia colorrectal en la muestra con el nivel de expresión medido de los dos o más genes biomarcadores de neoplasia colorrectal en un control, en el que una diferencia en el nivel de expresión medido de los dos o más genes en relación con el nivel de expresión medido de los dos o más genes en el control indica que el sujeto tiene o está en

riesgo de tener cáncer colorrectal; y seleccionar un plan clínico basado en que una diferencia en el nivel de expresión medido de los dos o más genes en relación con el nivel de expresión medido de los dos o más genes en el control indica que el sujeto tiene o está en riesgo de cáncer colorrectal.

Los métodos para determinar si un animal no humano está en riesgo de padecer un trastorno gastrointestinal pueden incluir las etapas de medir el nivel de expresión de uno o más genes de células B, células T o inmunoglobulinas en una muestra biológica del sujeto; comparar el nivel de expresión medido de uno o más genes de células B, células T o inmunoglobulinas en la muestra con el nivel de expresión medido de los uno o más genes de células B, células T o inmunoglobulinas en un control, en el que una diferencia en el nivel de expresión medido de uno o más genes en la muestra biológica en relación con el nivel de expresión medido de uno o más genes en el control indica que el sujeto está en riesgo de padecer un trastorno gastrointestinal. El trastorno gastrointestinal puede ser un linfoma gastrointestinal o una enfermedad inflamatoria intestinal. El animal no humano puede ser un gato o un perro. La muestra biológica puede ser una muestra de heces.

Breve descripción de los dibujos

Estas y otras características y ventajas de la invención presente serán divulgadas más completamente en, o considerada obvia por, la descripción detallada siguiente de la realización preferida de la invención, que va a ser considerada junto con los dibujos adjuntos en los que los números de referencia se refieren a partes similares y en los que además:

La Figura 1 es una ejecución de un archivo de electroforesis. El análisis electroforético se utilizó para comprobar la calidad del ARN extraído.

La Figura 2 es un electroferograma. El electroferograma se utilizó para comprobar la calidad del ARN extraído.

La Figura 3 es un resumen de la extracción. Esto representa un resumen del control de calidad de 120 muestras que se evaluaron utilizando el método de extracción descrito en el Ejemplo 2.

La Figura 4A es un gráfico que representa los resultados de un experimento en el que se comparan métodos de extracción de ARN total.

La Figura 4A muestra el número total de muestras que superan el control de calidad utilizando el método de extracción descrito en el Ejemplo 2.

La Figura 4B muestra cinco electroforesis realizadas con distintos métodos de extracción

La Figura 5A es un mapa de calor que muestra grupos de transcritos expresados diferencialmente. Como se muestra, se generó un mapa de calor utilizando 200 grupos de transcripción, que se asignaron a 187 genes diferentes, y las 265 muestras analizadas en el conjunto de entrenamiento. Las muestras están ordenadas por grupos: cáncer (rojo), adenoma precanceroso (naranja) y normal (verde).

La Figura 5B muestra los términos y vías GO expresados diferencialmente mediante paquete R de GAGE ($p < 0.05$). La altura de cada barra detalla el tamaño del conjunto de genes enriquecidos dentro de cada vía y la línea azul muestra el $-\log(p\text{-valor})$ para cada vía. La línea roja de puntos indica significancia ($p = 0.05$).

La figura 6 muestra un listado de los genes identificados como expresados diferencialmente en las neoplasias colorrectales.

La Figura 7A muestra una ejecución de electroforesis con 4 muestras felinas y 4 muestras caninas.

La Figura 7B muestra los Números de Integridad del ARN (RIN) para 4 muestras felinas individuales y 4 muestras caninas individuales y el promedio de las ocho muestras.

La Figura 7C muestra la concentración estimada de ARN eucariota (ng/uL) a partir de una ejecución de electroforesis, para 4 muestras felinas individuales y 4 muestras caninas individuales y la masa media para las 8 muestras.

La Figura 8A muestra los resultados de RT-qPCR para la región de la cadena pesada IgM C μ para identificar la presencia de transcritos relacionados con linfocitos en muestras caninas.

La Figura 8B muestra los resultados de RT-qPCR para dos recombinaciones de la región gamma del receptor de células T en muestras caninas.

La Figura 9 muestra los resultados de la RT-qPCR para Actina-B en muestras felinas.

La Figura 10A muestra 4 réplicas técnicas de 70,524 niveles de expresión de grupos de transcripción que utilizan arreglos de Transcriptoma Humano de Affymetrix.

La Figura 10B muestra 4 réplicas biológicas de 70,524 niveles de expresión de grupos de transcripción que utilizan arreglos de Transcriptoma Humano de Affymetrix.

La Figura 10C muestra 6 réplicas técnicas probadas con 6 meses de diferencia para analizar 70,524 niveles de expresión de grupos de transcripción que utilizan arreglos de Transcriptoma Humano de Affymetrix.

La Figura 11A muestra 4 réplicas técnicas de 5,149 niveles de expresión de grupos de transcripción que utilizan arreglos de Transcriptoma Humano de Affymetrix.

- 5 La Figura 11B muestra 4 réplicas biológicas de 5,149 niveles de expresión de grupos de transcripción que utilizan arreglos de Transcriptoma Humano de Affymetrix.

La Figura 11C muestra 6 réplicas técnicas probadas con 6 meses de diferencia para analizar 5,149 niveles de expresión de grupos de transcripción que utilizan arreglos de Transcriptoma Humano de Affymetrix.

- 10 La Figura 12A muestra un electroferograma comparativo y un archivo de electroforesis realizado a partir de métodos de extracción de la bibliografía.

La Figura 12B representa un electroferograma y un archivo de electroforesis ejecutados a partir de los métodos de extracción descritos en el Ejemplo 2.

La Figura 13 muestra un listado de genes identificados como expresados diferencialmente en neoplasias colorrectales, así como genes implicados en cáncer, neoplasias colorrectales y/o salud gastrointestinal.

- 15 La Figura 14 muestra un archivo de electroforesis ejecutado para 8 muestras derivadas de 4 gatos individuales que demuestra la consistencia de las firmas de ARN eucariotas y procariotas entre las réplicas biológicas.

La Figura 15A es una tabla que muestra los cebadores utilizados para el análisis RT-qPCR de muestras caninas.

La Figura 15B es una tabla que muestra los cebadores utilizados para el análisis RT-qPCR de muestras felinas.

- 20 La Figura 16A muestra los resultados de RT-qPCR para dos reordenamientos de inmunoglobulinas de células B para identificar la presencia de transcritos específicos de linfocitos en muestras caninas.

La Figura 16B muestra los resultados de RT-qPCR para los seis reordenamientos de la región gamma del receptor de células T para identificar la presencia de transcritos específicos de linfocitos en muestras felinas.

Descripción detallada

- 25 Esta descripción de las realizaciones preferidas debe leerse en relación con los dibujos adjuntos, que deben considerarse parte de la descripción escrita completa de esta invención. Las figuras de los dibujos no están necesariamente a escala y ciertas características de la invención pueden mostrarse a escala exagerada o de forma algo esquemática en aras de la claridad y la concisión. En la descripción, los términos relativos como "horizontal", "vertical", "arriba", "abajo", "superior" y "inferior", así como sus derivados (por ejemplo, "horizontalmente", "hacia abajo", "hacia arriba", etc.) deben interpretarse como referidos a la orientación descrita o mostrada en la figura del dibujo en cuestión. Estos términos relativos se utilizan para facilitar la descripción y normalmente no pretenden exigir una orientación determinada. Los términos "hacia dentro" frente a "hacia fuera", "longitudinal" frente a "lateral" y similares deben interpretarse entre sí o en relación con un eje de elongación, o un eje o centro de rotación, según proceda. Los términos relativos a fijaciones, acoplamientos y similares, tales como "conectado" e "interconectado", se refieren a una relación en la que las estructuras están fijadas o unidas entre sí, ya sea directa o indirectamente a través de estructuras intermedias, así como a fijaciones o relaciones tanto móviles como rígidas, a menos que se describa expresamente lo contrario. El término "conectado operativamente" se refiere a un accesorio, acoplamiento o conexión que permite que las estructuras pertinentes funcionen de acuerdo con lo previsto en virtud de esa relación. Cuando se ilustre una sola máquina, el término "máquina" también incluirá cualquier conjunto de máquinas que, individual o conjuntamente, ejecuten un conjunto (o varios conjuntos) de instrucciones para realizar una cualquiera o más de las metodologías descritas en el presente documento.
- 30
- 35
- 40

- La presente invención se basa en parte en el desarrollo por parte de los inventores de un método para separar células eucariotas de células bacterianas en una muestra de heces, por ejemplo, una muestra de heces obtenida de un mamífero. En el colon, hay aproximadamente 10^{12} células bacterianas por gramo de contenido intestinal. Esta microflora colónica puede incluir entre 300-1000 especies. Una muestra de heces o materia fecal es una mezcla macromolecular compleja que incluye no sólo células eucariotas desprendidas del lumen intestinal del tracto gastrointestinal, sino también microbios, incluidas bacterias y cualquier parásito gastrointestinal, residuos de alimentos no digeribles ni absorbidos, secreciones de células intestinales y material excretado, como mucosa y pigmentos. Las heces normales están compuestas por un 75 % de agua y un 25 % de materia sólida. Las bacterias constituyen aproximadamente el 60 % de la masa seca total de las heces. La elevada carga bacteriana puede contribuir a una relación señal-ruido desfavorable para la detección de biomarcadores eucarióticos a partir de una muestra de heces. Además, las señales eucariotas pueden degradarse mucho. La extracción y el procesamiento de estos ácidos nucleicos eucariotas pueden favorecer o acelerar su degradación, lo que limita gravemente los análisis posteriores.
- 45
- 50

Los métodos aquí divulgados incluyen métodos para aislar ácidos nucleicos eucariotas de una muestra de heces. Dichos ácidos nucleicos eucariotas pueden evaluarse para determinar los niveles de biomarcadores específicos que

pueden ser indicativos de un trastorno o enfermedad gastrointestinal en una eucariota, por ejemplo, un mamífero. El mamífero puede ser un ser humano o un animal no humano, por ejemplo, un ser humano, un perro, un gato, un primate no humano, un rumiante, un úrsido, un équido, un cerdo, una oveja, una cabra, un camélido, un búfalo, un ciervo, un alce americano, un alce, un mustélido, un conejo, una cobaya, un hámster, una rata, un ratón, un paquidermo, un rinoceronte o una chinchilla.

Los inventores han encontrado que podían separar eficazmente las células eucariotas de las células bacterianas en una muestra de heces eucariotas. Los inventores también han encontrado que podían detectar biomarcadores eucarióticos en el ARN aislado de dichas células eucarióticas. Dichos biomarcadores pueden ser útiles para la detección de trastornos gastrointestinales, por ejemplo, cáncer colorrectal, enfermedad celíaca, enfermedad de Crohn, gastritis, gastroenteritis, cáncer gástrico, úlceras gástricas, enterocolitis necrotizante, tumores del estroma gastrointestinal, linfoma gastrointestinal, neoplasia gastrointestinal, linfoma sarcoma, adenocarcinoma, enfermedad inflamatoria intestinal, síndrome del intestino irritable, neoplasia pancreática, neoplasia hepática, colangiocarcinoma, colitis, virus de la leucemia felina, diarrea vírica bovina, síndrome de hemorragia yeyunal, gastroenteritis, fiebre catarral maligna, panleucopenia felina, fibrosis del intestino delgado, enfermedad infiltrativa del colon, criptosporidiosis, coccidiosis y otras infecciones parasitarias zoonóticas, o infecciones por coronavirus, parvovirus, astrovirus, norovirus o rotavirus. Se proporcionan aquí materiales y métodos para determinar si un sujeto, por ejemplo, un humano, un perro o un gato, tiene riesgo de padecer una enfermedad gastrointestinal, por ejemplo, cáncer colorrectal, linfoma o enfermedad inflamatoria intestinal u otras enfermedades. También se proporcionan métodos para el diagnóstico de enfermedades y métodos para identificar el estado de salud de un sujeto.

Los métodos aquí divulgados son general y diversamente útiles para la detección, diagnóstico y tratamiento de la salud gastrointestinal. Los métodos de detección pueden incluir la medición del nivel de expresión en una muestra de heces de uno, dos o más biomarcadores en una muestra de un sujeto, por ejemplo, un paciente, que tiene un trastorno gastrointestinal o se sospecha que tiene un trastorno gastrointestinal y la comparación del nivel de expresión medido con el nivel de expresión medido de uno, dos o más biomarcadores en un control. Una diferencia en el nivel de expresión medido de uno, dos o más biomarcadores en la muestra de un sujeto en relación con el nivel de expresión medido de uno, dos o más biomarcadores en un control es una indicación de que el sujeto padece un trastorno gastrointestinal. En algunas realizaciones, una diferencia en el nivel de expresión medido de uno, dos o más biomarcadores en la muestra de un sujeto en relación con el nivel de expresión medido de uno, dos o más biomarcadores en un control es una indicación de que el sujeto, por ejemplo, un paciente, está en riesgo de padecer un trastorno gastrointestinal.

En otra realización, los métodos de detección de la enfermedad pueden incluir la medición de la proporción del nivel de expresión relativa, por ejemplo, las proporciones relativas, de uno, dos o más biomarcadores en la muestra de heces de un sujeto y la comparación de la proporción relativa de estos biomarcadores con la proporción del nivel de expresión relativa de uno, dos o más biomarcadores en un control. Una diferencia en la proporción del nivel de expresión relativa medida de uno, dos o más biomarcadores en la muestra de un sujeto en relación con un control es una indicación de que el sujeto padece una enfermedad gastrointestinal. En algunas realizaciones, una diferencia en la proporción del nivel de expresión medido de uno, dos o más biomarcadores en la muestra de un sujeto en relación con la proporción del nivel de expresión medido de uno, dos o más biomarcadores en un control es una indicación de que el sujeto está en riesgo de padecer un trastorno gastrointestinal.

Los métodos pueden incluir determinar el nivel de expresión de dos o más biomarcadores de neoplasia colorrectal en el ARN humano aislado de una muestra de heces obtenida de un paciente al determinar si los niveles de los dos o más genes biomarcadores de neoplasia colorrectal son diferentes en relación con los niveles de los mismos dos o más genes biomarcadores de neoplasia colorrectal en un control. En la Figura 6 (Panel A) y la Figura 13 (Panel B) se muestran genes biomarcadores de neoplasias colorrectales ejemplares. Algunos o todos los genes biomarcadores de neoplasias colorrectales listados en la Figura 6 (Panel A) y en la Figura 13 (Panel B) pueden formar un panel. En algunas realizaciones, los genes biomarcadores de neoplasia colorrectal listados en la Figura 6 (Panel A) y la Figura 13 (Panel B) también pueden incluir subconjuntos de genes biomarcadores de neoplasia colorrectal. Las composiciones pueden incluir arreglos de genes y conjuntos de sondas configurados para la detección específica de los paneles de marcadores aquí divulgados. Las composiciones aquí descritas también pueden incluir kits que comprenden arreglos de genes y conjuntos de sondas configurados para la detección específica de los paneles de marcadores aquí divulgados.

Los genes biomarcadores de neoplasias colorrectales y los paneles de genes biomarcadores de neoplasias colorrectales pueden utilizarse en el diagnóstico del cáncer colorrectal o de lesiones precancerosas. Un biomarcador es generalmente una característica que puede medirse y cuantificarse objetivamente y utilizarse para evaluar un proceso biológico, por ejemplo, el desarrollo, la progresión, la remisión y la recurrencia de una neoplasia colorrectal. Los biomarcadores pueden adoptar muchas formas, como ácidos nucleicos, polipéptidos, metabolitos o parámetros físicos o fisiológicos.

Estos biomarcadores de células eucariotas pueden incluir: a) una secuencia de ácido desoxirribonucleico (ADN), b) una secuencia de ácido ribonucleico (ARN), c) una secuencia prevista de aminoácidos, que comprenden la columna vertebral de la proteína, d) niveles de expresión de biomarcadores de ácido ribonucleico, e) un nivel de expresión previsto de una secuencia de aminoácidos o f) cualquier combinación de los anteriores. En algunas realizaciones, el

biomarcador puede ser un marcador de células T o un marcador de células B. En algunas realizaciones, el biomarcador puede ser un biomarcador útil para la detección de la expansión clonal de células T o células B. Los biomarcadores ejemplares pueden incluir IgA, IgM, IgG, IgE, IgD, el receptor gamma de células T, el receptor alfa de células T, el receptor beta de células T, la región del receptor delta de células T o la región determinante complementaria de células B. En algunas realizaciones, se pueden utilizar biomarcadores para la normalización, como el GADPH. En algunas realizaciones, los biomarcadores pueden utilizarse para la detección de tipos celulares específicos, tal como el uso de Actina-B para la detección de células epiteliales o IgM C-mu para la detección de linfocitos. En algunas realizaciones, los biomarcadores pueden utilizarse para la detección de marcadores específicos de la enfermedad, como la detección del receptor gamma de células T para la detección de la expansión clonal de linfocitos. En algunas realizaciones, los biomarcadores pueden comprender los de detección de virus y parásitos, incluidos los de cápsides, capsómeros, replicasas y proteínas de la pared del ooqueste.

Una muestra biológica descrita en el presente documento puede ser una muestra que contiene células u otro material celular del que se pueden obtener ácidos nucleicos u otros analitos. Una muestra biológica puede ser una muestra de control o una muestra experimental. En la presente invención, una muestra biológica es una muestra de heces. La muestra biológica puede obtenerse inmediatamente después de la defecación en un retrete, en el suelo o en un dispositivo de recogida. En algunas realizaciones, la muestra biológica puede obtenerse tras un procedimiento, tal como un enema o una endoscopia. La muestra biológica puede analizarse inmediatamente. Alternativamente, la muestra biológica puede almacenarse en un tampón antes de la prueba, por ejemplo un tampón acuoso, un tampón a base de glicerol, un tampón a base de disolvente polar, un tampón de equilibrio osmótico u otro tampón suficiente para preservar la muestra biológica. Además, o alternativamente, la muestra biológica puede recogerse y almacenarse refrigerada, por ejemplo, a 4 °C, o congelada, por ejemplo, a 0 °C, -20 °C, -80 °C, -140 °C, o menos antes de la prueba. La muestra biológica puede almacenarse durante 1 mes, 2 meses, 4 meses, 6 meses, 1 año, 2 años o más antes de la prueba.

La muestra biológica puede proceder de una eucariota, por ejemplo un mamífero. El mamífero puede ser un ser humano o un animal no humano, por ejemplo, un ser humano, un perro, un gato, un primate no humano, un rumiante, un úrsido, un équido, un cerdo, una oveja, una cabra, un camélido, un búfalo, un ciervo, un alce americano, un alce, un mustélido, un conejo, una cobaya, un hámster, una rata, un ratón, un paquidermo, un rinoceronte o una chinchilla.

Métodos

En el presente documento se proporcionan métodos útiles para el aislamiento de ácidos nucleicos a partir de una muestra de heces, que están enriquecidos para ácidos nucleicos eucariotas. El método puede incluir la disolución de la muestra de heces con tampón. La muestra puede someterse a agitación de vórtex, zarandeo, agitación, rotación u otro método de agitación suficiente para dispersar los sólidos y las bacterias fecales. La temperatura a la que se llevan a cabo las etapas de agitación y centrifugación puede variar, por ejemplo, de aproximadamente 4 °C a aproximadamente 20 °C, de aproximadamente 4 °C a aproximadamente 1 °C, de aproximadamente 4 °C a aproximadamente 10 °C, de aproximadamente 4 °C a aproximadamente 6 °C. Tras la disrupción, la muestra puede someterse a una o más rondas de centrifugación. En algunas realizaciones, la etapa de disrupción y la centrifugación pueden repetirse una, dos, tres o más veces adicionales. Los reactivos disponibles comercialmente, por ejemplo los reactivos Nuclisens® EasyMag® pueden utilizarse para la disrupción fecal, el lavado y la lisis celular. El tampón de lisis también puede utilizarse para lisar las células eucariotas. El lisado puede centrifugarse más y el sobrenadante utilizarse para introducirlo en una máquina automatizada de aislamiento de ARN, por ejemplo el instrumento EasyMag®. En algunas realizaciones, los ácidos nucleicos extraídos pueden tratarse con DNasa para limpiar la solución de ADN. Pueden utilizarse otros métodos, como la disrupción celular mecánica o enzimática seguida de un método en fase sólida como la cromatografía en columna o la extracción con disolventes orgánicos, por ejemplo, la extracción con fenol-cloroformo o tiocianato-fenol-cloroformo. En algunas realizaciones, el ácido nucleico puede extraerse en una microesfera funcionalizada. En algunas realizaciones, la perla funcionalizada puede comprender además un núcleo magnético ("perla magnética"). En algunas realizaciones, la perla funcionalizada puede incluir una superficie funcionalizada con una fracción cargada. La fracción cargada puede seleccionarse entre: amina, ácido carboxílico, carboxilato, amina cuaternaria, sulfato, sulfonato o fosfato.

Para la extracción de ácidos nucleicos, la muestra de heces se disgrega en presencia de uno o más de un tampón, un tensioactivo y un inhibidor de ribonucleasas para formar una suspensión. El tampón puede ser un tampón biológicamente compatible, por ejemplo, solución salina equilibrada de Hanks, solución de Alsever, solución salina equilibrada de Earle, solución salina equilibrada de Gey, solución salina tamponada con fosfato, solución salina equilibrada de Puck, solución salina equilibrada de Ringer, solución salina equilibrada de Simm, solución salina tamponada con TRIS o solución salina equilibrada de Tyrode. El tensioactivo puede ser un tensioactivo iónico o no iónico, por ejemplo, Tween-20, o Triton-X-100. El inhibidor de ribonucleasas puede estar basado en disolventes, proteínas u otro tipo de método para evitar la destrucción del ARN, incluyendo, por ejemplo, Inhibidor Protector de RNasa (Roche), RNasin® (Promega), SUPERase-In™ (Thermo Fisher Scientific), RNaseOUT™ (Thermo Fisher Scientific), ANTI-RNasa, Inhibidor Recombinante RNasa o un inhibidor de RNasa clonado. La muestra de heces puede disolverse de diversas maneras, por ejemplo mediante agitación de vórtex, zarandeo, agitación, rotación u otro método de agitación suficiente para dispersar los sólidos y las bacterias fecales. En algunas realizaciones, la muestra de heces puede disolverse utilizando: perlas recubiertas, perlas magnéticas o un instrumento de agitación, como una varilla de vidrio, una varilla metálica, un palo de madera o una cuchilla de madera.

A continuación, la suspensión se separa en una porción líquida y una porción sólida. La separación puede llevarse a cabo, por ejemplo, mediante centrifugación, filtración, sondas diana que se unen específicamente a células eucariotas, anticuerpos, filtración en columna, filtración en perlas o métodos cromatográficos. La porción líquida está enriquecida para ácido nucleico bacteriano y puede desecharse. La porción sólida puede resuspenderse en un tampón en presencia o ausencia de un tensioactivo y en presencia o ausencia de una ribonucleasa. La etapa de separación puede repetirse una, dos, tres, cuatro, cinco, seis, siete, ocho o más veces.

La temperatura a la que se llevan a cabo las etapas de disrupción y separación puede variar, por ejemplo, de aproximadamente 4 °C a aproximadamente 20 °C, de aproximadamente 4 °C a aproximadamente 15 °C, de aproximadamente 4 °C a aproximadamente 10 °C, de aproximadamente 4 °C a aproximadamente 6 °C.

El gránulo resultante obtenido de la etapa de separación se suspende en un tampón de lisis, por ejemplo, un tampón que comprende un agente caotrópico y opcionalmente un tensioactivo para formar un lisado. En algunas realizaciones, el agente caotrópico puede ser tiocianato de guanidio y el tensioactivo puede ser Triton-X-100. En algunas realizaciones, el tampón de lisis puede incluir o excluir Tris-HCl, ácido etilendiaminotetraacético (EDTA), dodecilsulfato sódico (SDS), Nonidet P-40, desoxicolato sódico o ditiotreitól.

El lisado se fracciona en una porción enriquecida para ácido nucleico eucariota. El fraccionamiento puede realizarse, por ejemplo, mediante centrifugación, filtración, sondas diana que se unen específicamente al ácido nucleico eucariota, anticuerpos, filtración en columna, filtración en perlas o métodos cromatográficos. El fraccionamiento por centrifugación puede dar lugar a la formación de una capa inferior (un gránulo), que comprende restos celulares, una capa media hidrófila que comprende ácidos nucleicos eucariotas, y una capa superior hidrófoba que comprende lípidos y fracciones de membrana. La capa intermedia se puede recoger. En algunas realizaciones, la capa intermedia y la capa superior pueden recogerse juntas. La capa intermedia puede recogerse a través de un orificio de paso estrecho. El orificio de paso estrecho puede ser una punta de pipeta o una jeringa provista de una aguja. La punta de pipeta puede ser, por ejemplo, de 1, 5, 10, 20 ul o 100 ul. La aguja puede ser, por ejemplo, de calibre 18 o de calibre 15.

La capa recogida que contiene ácidos nucleicos eucariotas puede someterse a extracción adicional. El método de extracción posterior puede variar. Los métodos ejemplares incluyen métodos basados en partículas magnéticas, métodos basados en columnas, métodos basados en filtros, métodos basados en perlas o métodos basados en disolventes orgánicos. Los métodos basados en partículas magnéticas pueden incluir reactivos disponibles en el mercado, por ejemplo los reactivos Nuclisens® EasyMag® (bioMerieux).

Los ácidos nucleicos extraídos pueden analizarse en busca de biomarcadores eucariotas que sean relevantes para los trastornos gastrointestinales o las células gastrointestinales. Los biomarcadores pueden proporcionar información sobre la salud de un individuo, es decir, del sujeto. Estos biomarcadores de células eucariotas pueden incluir: a) una secuencia de ácido desoxirribonucleico (ADN), b) una secuencia de ácido ribonucleico (ARN), c) una secuencia prevista de aminoácidos, que comprenden la columna vertebral de las proteínas, d) niveles de expresión o proporciones de niveles de expresión de biomarcadores de ARN, e) un nivel de expresión previsto o una proporción de nivel de expresión previsto de una secuencia de aminoácidos, o f) cualquier combinación de las anteriores. El aislamiento de biomarcadores a partir de células eucariotas puede permitir la comparación entre una muestra experimental y un control. El aislamiento de estos biomarcadores a partir de células eucariotas puede proporcionar un método de detección de la enfermedad intestinal en la muestra experimental. La comparación puede incluir la evaluación de: a) la variación en una secuencia de ADN, b) la variación en una secuencia de ARN, c) la variación en la secuencia de aminoácidos predicha, d) la variación en los niveles de expresión o la variación de la proporción de los niveles de expresión de los biomarcadores de ARN, e) la variación en el nivel de expresión predicho o la variación en la proporción del nivel de expresión predicho de una secuencia de aminoácidos, o f) una variación que constituya cualquier combinación de las anteriores. Se puede determinar una variación cuando el biomarcador medido de una muestra experimental es diferente del biomarcador medido en un control.

El método puede incluir la obtención de una muestra experimental y un control, por ejemplo, una muestra de heces. La muestra de heces contiene células eucariotas desprendidas que pueden evaluarse en busca de biomarcadores. En algunas realizaciones, las células eucariotas pueden ser enterocitos, linfocitos, células similares a la enterocromafina, células enteroendocrinas, células neuroendocrinas, células pancreáticas, células hepáticas, células gástricas u otras. El método permite evaluar las células eucariotas de la muestra de heces para detectar biomarcadores eucariotas. Los biomarcadores pueden incluir una secuencia de ADN, una secuencia de ARN, una secuencia prevista de aminoácidos, un nivel de expresión o proporción de nivel de expresión de biomarcadores de ARN, un nivel de expresión previsto o una proporción de nivel de expresión previsto de una secuencia de aminoácidos o cualquier combinación de los anteriores. En un aspecto, la etapa de evaluación comprende cualquier tipo de secuenciación de microarreglos, reacción en cadena de la polimerasa (PCR), secuenciación de ácidos nucleicos, código de barras molecular o captura de sondas.

Los métodos de la invención y las composiciones aquí descritas también son útiles para seleccionar un plan clínico para un individuo que sufre una enfermedad intestinal. Mediante este método, el plan clínico puede incluir la administración de otros procedimientos diagnósticos. En algunas realizaciones, el plan clínico puede incluir un método de tratamiento.

Los niveles de los biomarcadores eucarióticos pueden evaluarse mediante diversos métodos. Los niveles de expresión pueden determinarse a nivel de ácido nucleico, por ejemplo, a nivel de ARN, o a nivel de polipéptido. La expresión de ARN puede abarcar la expresión de ARN total, ARNm, ARNt, ARNr, ARNnc, ARNsm, miARN y snoARN. La expresión a nivel de ARN puede medirse directa o indirectamente midiendo los niveles de ADNc correspondientes al ARN pertinente. Como alternativa, o además, también pueden analizarse los polipéptidos codificados por el ARN, los reguladores de ARN de los genes que codifican los factores de transcripción pertinentes, y los niveles de los polipéptidos del factor de transcripción. Los métodos para determinar la expresión génica a nivel del ARNm incluyen, por ejemplo, el análisis de microarreglos, el análisis en serie de la expresión génica (SAGE), la RT-PCR, la secuenciación, la hibridación basada en ensayos de cuantificación de códigos de barras digitales, la RT-PCR multiplex, la PCR de gota digital (ddPCR), los espectrofotómetros NanoDrop, la RT-qPCR, la qPCR, la espectroscopia UV, la secuenciación del ARN, la secuenciación de próxima generación, los ensayos de hibridación basados en lisados que utilizan amplificación de señales de ADN ramificado, tal como el QuantiGene 2.0 Single Plex, y los métodos de análisis de ADN ramificado Single Plex, y métodos de análisis de ADN ramificado. Los ensayos digitales de cuantificación de códigos de barras pueden incluir el BeadArray (Illumina), los sistemas xMAP (Luminex), el nCounter (Nanostring), el molecular High Throughput Genomics (HTG), el BioMark (Fluidigm) o el microarreglo Wafergen. Los ensayos pueden incluir DASL (Illumina), RNA-Seq (Illumina), TruSeq (Illumina), SureSelect (Agilent), Bioanalyzer (Agilent) y TaqMan (ThermoFisher).

Podemos utilizar los términos “ácido nucleico” y “polinucleótido” indistintamente para referirnos tanto al ARN como al ADN, incluidos el ADNc, el ADN genómico, el ADN sintético y el ADN (o ARN) que contiene análogos de ácido nucleico, cualquiera de los cuales puede codificar un polipéptido de la invención y todos los cuales están comprendidos en la invención. Los polinucleótidos pueden tener esencialmente cualquier estructura tridimensional. Un ácido nucleico puede ser de cadena doble o de cadena sencilla (es decir, una cadena sentido o una cadena antisentido). Ejemplos no limitantes de polinucleótidos incluyen genes, fragmentos de genes, exones, intrones, ARN mensajero (ARNm) y porciones del mismo, ARN de transferencia, micro ARN, ARN ribosómico, ARNsí, micro ARN, ribozimas, ADNc, polinucleótidos recombinantes, polinucleótidos ramificados, plásmidos, vectores, ADN aislado de cualquier secuencia, ARN aislado de cualquier secuencia, sondas de ácidos nucleicos y cebadores, así como análogos de ácidos nucleicos. En el contexto de la presente invención, los ácidos nucleicos pueden codificar un fragmento de un biomarcador, por ejemplo, un biomarcador de expansión clonal de células B o células T, o una variante biológicamente activa del mismo.

Un ácido nucleico “aislado” puede ser, por ejemplo, una molécula de ADN o un fragmento de la misma, siempre que al menos una de las secuencias de ácido nucleico que normalmente se encuentran inmediatamente flanqueando a esa molécula de ADN en un genoma esté eliminada o ausente. Así, un ácido nucleico aislado incluye, sin limitación, una molécula de ADN que existe como molécula separada, independiente de otras secuencias (por ejemplo, un ácido nucleico sintetizado químicamente, o un fragmento de ADNc o ADN genómico producido por la reacción en cadena de la polimerasa (PCR) o el tratamiento con endonucleasas de restricción). Un ácido nucleico aislado también se refiere a una molécula de ADN que se incorpora a un vector, a un plásmido de replicación autónoma, a un virus o al ADN genómico de un procarionte o eucariote. Además, un ácido nucleico aislado puede incluir un ácido nucleico manipulado, tal como una molécula de ADN que forma parte de un ácido nucleico híbrido o de fusión. Un ácido nucleico que existe entre muchos (por ejemplo, docenas, o cientos a millones) de otros ácidos nucleicos dentro de, por ejemplo, bibliotecas de ADNc o bibliotecas genómicas, o cortes de gel que contienen una digestión de restricción de ADN genómico, no es un ácido nucleico aislado.

Las moléculas de ácido nucleico aisladas pueden producirse de diversas maneras. Por ejemplo, pueden utilizarse técnicas de reacción en cadena de la polimerasa (PCR) para obtener un ácido nucleico aislado que contenga una secuencia de nucleótidos descrita en el presente documento, incluidas las secuencias de nucleótidos que codifican un polipéptido descrito en el presente documento. La PCR puede utilizarse para amplificar secuencias específicas tanto de ADN como de ARN, incluidas secuencias de ADN genómico total o de ARN celular total. Generalmente, la información de la secuencia de los extremos de la región de interés o más allá se emplea para diseñar cebadores oligonucleotídicos que son idénticos o similares en secuencia a las cadenas opuestas de la plantilla a amplificar. También se dispone de varias estrategias de PCR mediante las cuales se pueden introducir modificaciones específicas de la secuencia de nucleótidos en un ácido nucleico molde.

Los ácidos nucleicos aislados también pueden ser sintetizados químicamente, ya sea como una sola molécula de ácido nucleico (por ejemplo, utilizando la síntesis automatizada de ADN en la dirección 3' a 5' utilizando la tecnología de fosforamidita) o como una serie de oligonucleótidos. Por ejemplo, se pueden sintetizar uno o más pares de oligonucleótidos largos (por ejemplo, >50-100 nucleótidos) que contengan la secuencia deseada, y cada par debe contener un segmento corto de complementariedad (por ejemplo, aproximadamente 15 nucleótidos) de manera que se forme un dúplex cuando el par de oligonucleótidos se hibrida. La ADN polimerasa se utiliza para extender los oligonucleótidos, lo que da lugar a una única molécula de ácido nucleico de doble cadena por par de oligonucleótidos, que puede ligarse a un vector.

Dos ácidos nucleicos o los polipéptidos que codifican pueden describirse como poseedores de un cierto grado de identidad entre sí. Por ejemplo, un gen biomarcador de neoplasia colorrectal seleccionado de la Figura 6 (Panel A) o de la Figura 13 (Panel B) y una variante biológicamente activa del mismo pueden describirse como exhibiendo un cierto grado de identidad. Los alineamientos pueden ensamblarse localizando secuencias cortas en el sitio Protein

Information Research (PIR) (<http://pir.georgetown.edu>), seguido de un análisis con el algoritmo “secuencias cortas casi idénticas” Basic Local Alignment Search Tool (BLAST) en el sitio web del NCBI (<http://www.ncbi.nlm.nih.gov/blast>).

Tal como se utiliza en el presente documento, el término “porcentaje de identidad de secuencia” se refiere al grado de identidad entre cualquier secuencia de consulta dada y una secuencia sujeto. Por ejemplo, una secuencia génica de biomarcador de neoplasia colorrectal listada en la Figura 6 (Panel A) o en la Figura 13 (Panel B) puede ser la secuencia de consulta y un fragmento de una secuencia génica de biomarcador de neoplasia colorrectal listada en la Figura 6 (Panel A) o en la Figura 13 (Panel B) puede ser la secuencia objeto. Del mismo modo, un fragmento de una secuencia génica de biomarcador de neoplasia colorrectal listada en la Figura 6 (Panel A) o en la Figura 13 (Panel B) puede ser la secuencia de consulta y una variante biológicamente activa de la misma puede ser la secuencia objeto.

Para determinar la identidad de secuencia, una secuencia de ácido nucleico o aminoácido de consulta puede alinearse con una o más secuencias de ácido nucleico o aminoácido del sujeto, respectivamente, utilizando un programa informático, por ejemplo, ClustalW (versión 1.83, parámetros por defecto), que permite realizar alineaciones de secuencias de ácidos nucleicos o proteínas en toda su longitud (alineación global).

Los ácidos nucleicos y polipéptidos aquí descritos pueden denominarse “exógenos”. El término “exógeno” indica que el ácido nucleico o el polipéptido forma parte de, o está codificado por, una construcción de ácido nucleico recombinante, o no se encuentra en su entorno natural. Por ejemplo, un ácido nucleico exógeno puede ser una secuencia de una especie introducida en otra especie, es decir, un ácido nucleico heterólogo. Normalmente, dicho ácido nucleico exógeno se introduce en la otra especie a través de una construcción de ácido nucleico recombinante. Un ácido nucleico exógeno también puede ser una secuencia que es nativa de un organismo y que ha sido reintroducida en células de ese organismo. Un ácido nucleico exógeno que incluye una secuencia nativa a menudo puede distinguirse de la secuencia nativa por la presencia de secuencias no naturales ligadas al ácido nucleico exógeno, por ejemplo, secuencias reguladoras no nativas que flanquean una secuencia nativa en una construcción de ácido nucleico recombinante. Además, los ácidos nucleicos exógenos transformados de forma estable suelen integrarse en posiciones distintas a la posición en la que se encuentra la secuencia nativa.

Los ácidos nucleicos pueden incluir ácidos nucleicos que tienen una secuencia de nucleótidos de uno cualquiera de los biomarcadores de neoplasia colorrectal listados en la Figura 6 (Panel A) o la Figura 13 (Panel B), o una secuencia de ácido nucleico que es al menos aproximadamente 70 %, al menos aproximadamente 75 %, al menos aproximadamente 80 %, al menos aproximadamente 85 %, al menos aproximadamente 90 %, al menos aproximadamente 95 %, al menos aproximadamente 99 % idéntica a los ácidos nucleicos listados en la Figura 6 (Panel A) o la Figura 13 (Panel B).

Un ácido nucleico, por ejemplo, un oligonucleótido (por ejemplo, una sonda o un cebador) que es específico para un ácido nucleico diana se hibridará con el ácido nucleico diana en condiciones adecuadas. Podemos referirnos a la hibridación o al proceso de hibridación como el proceso por el cual una cadena simple de oligonucleótidos se hibrida a una cadena complementaria mediante el apareamiento de bases en condiciones de hibridación definidas. Se trata de una interacción específica, es decir, no aleatoria, entre dos polinucleótidos complementarios. La hibridación y la fuerza de la hibridación (es decir, la fuerza de la asociación entre los ácidos nucleicos) están influidas por factores como el grado de complementariedad entre los ácidos nucleicos, la rigurosidad de las condiciones implicadas y la temperatura de fusión (T_m) del híbrido formado. Los productos de hibridación pueden ser dúplex o triplex formados con dianas en solución o sobre soportes sólidos.

En algunas realizaciones, los ácidos nucleicos pueden incluir secuencias cortas de ácido nucleico útiles para el análisis y la cuantificación de los genes biomarcadores de neoplasia colorrectal listados en la Figura 6 o 13. Dichos ácidos nucleicos aislados pueden ser cebadores oligonucleotídicos. En general, un cebador oligonucleotídico es un oligonucleótido complementario a una secuencia de nucleótidos diana, por ejemplo, la secuencia de nucleótidos de cualquiera de los genes biomarcadores de neoplasias colorrectales listados en la Figura 6 o 13, que puede servir como punto de partida para la síntesis de ADN mediante la adición de nucleótidos al extremo 3' del cebador en presencia de una ADN o ARN polimerasa. El nucleótido 3' del cebador debe ser generalmente idéntico a la secuencia diana en una posición nucleotídica correspondiente para una extensión y/o amplificación óptima. Los cebadores pueden adoptar muchas formas, incluyendo, por ejemplo, cebadores de ácido nucleico peptídico, cebadores de ácido nucleico bloqueados, cebadores de ácido nucleico desbloqueados y/o cebadores modificados con fosforotioato. En algunas realizaciones, un cebador directo puede ser un cebador que es complementario a la cadena antisentido del dsADN y un cebador inverso puede ser un cebador complementario a la cadena sentido del dsADN. También podemos referirnos a pares de cebadores. En algunas realizaciones, un par de cebadores 5' diana puede ser un par de cebadores que incluye al menos un cebador directo y al menos un cebador inverso que amplifica la región 5' de una secuencia nucleotídica diana. En algunas realizaciones, un par de cebadores diana 3' puede ser un par de cebadores con al menos un cebador directo y al menos un cebador inverso que amplifica la región 3' de una secuencia nucleotídica diana. En algunas realizaciones, el cebador puede incluir un marcado detectable, como se explica a continuación. En algunas realizaciones, el marcado detectable puede ser un marcado cuantificable.

Los cebadores de oligonucleótidos son útiles para la amplificación de cualquiera de las secuencias de genes biomarcadores de neoplasias colorrectales listadas en la Figura 6 (Panel A) o en la Figura 13 (Panel B). En algunas realizaciones, los cebadores oligonucleotídicos pueden ser complementarios a dos o más de los genes biomarcadores

de neoplasia colorrectal divulgados en el presente documento, por ejemplo, los genes biomarcadores de neoplasia colorrectal listados en la Figura 6 (Panel A) o la Figura 13 (Panel B). La longitud del cebador puede variar en función de la secuencia de bases de nucleótidos y de la composición de la secuencia de ácido nucleico concreta de la sonda y del método específico para el que se utilice la sonda. En general, las longitudes útiles de los cebadores pueden ser de aproximadamente 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 bases de nucleótido. Las longitudes útiles de los cebadores pueden oscilar desde 8 bases de nucleótidos a aproximadamente 60 bases de nucleótidos; desde aproximadamente 12 bases de nucleótidos a aproximadamente 50 bases de nucleótidos; desde aproximadamente 12 bases de nucleótidos a aproximadamente 45 bases de nucleótidos; desde aproximadamente 12 bases de nucleótidos a aproximadamente 40 bases de nucleótidos; desde aproximadamente 12 bases de nucleótidos a aproximadamente 35 bases de nucleótidos; desde aproximadamente 15 bases de nucleótidos a aproximadamente 40 bases de nucleótidos; desde aproximadamente 15 bases de nucleótidos a aproximadamente 35 bases de nucleótidos; desde aproximadamente 18 bases de nucleótidos a aproximadamente 50 bases de nucleótidos; desde aproximadamente 18 bases de nucleótidos a aproximadamente 40 bases de nucleótidos; desde aproximadamente 18 bases de nucleótidos a aproximadamente 35 bases de nucleótidos; desde aproximadamente 18 bases de nucleótidos a aproximadamente 30 bases de nucleótidos; desde aproximadamente 20 bases de nucleótidos a aproximadamente 30 bases de nucleótidos; desde aproximadamente 20 bases de nucleótidos a aproximadamente 25 bases de nucleótidos.

Las sondas, es decir, los fragmentos aislados de ácido nucleico que se unen selectivamente y son complementarios a cualquiera de las secuencias de genes biomarcadores de neoplasias colorrectales listadas en la Figura 6 (Panel A) y la Figura 13 (Panel B) pueden ser oligonucleótidos o polinucleótidos, ADN o ARN, de cadena simple o doble, y naturales o modificados, ya sea en las bases de nucleótidos o en la columna vertebral. Las sondas pueden producirse por diversos métodos, como la síntesis química o enzimática.

La longitud de la sonda puede variar en función de la secuencia de bases de nucleótidos y la composición de la secuencia de ácido nucleico particular de la sonda y el método específico para el que se utiliza la sonda. En general, las longitudes de sonda útiles pueden ser de aproximadamente 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 50, 55, 60, 65, 70, 75, 80, 85, 90, 100, 110, 120, 140, 150, 175 o 200 bases de nucleótidos. En general, las longitudes de sondas útiles oscilarán desde aproximadamente 8 a aproximadamente 200 bases de nucleótidos; desde aproximadamente 12 a aproximadamente 175 bases de nucleótidos; desde aproximadamente 15 a aproximadamente 150 bases de nucleótidos; desde aproximadamente 15 a aproximadamente 100 bases de nucleótidos; desde aproximadamente 15 a aproximadamente 75 bases de nucleótidos; desde aproximadamente 15 a aproximadamente 60 bases de nucleótidos; desde aproximadamente 20 a aproximadamente 100 bases de nucleótidos; desde aproximadamente 20 a aproximadamente 75 bases de nucleótidos; desde aproximadamente 20 a aproximadamente 60 bases de nucleótidos; desde aproximadamente 20 a aproximadamente 50 bases de nucleótidos de longitud. En algunas realizaciones el conjunto de sondas puede comprender sondas dirigidas a al menos 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 40, 45, 50, 60, 70, 80, 90, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575 o más, o todos, los genes biomarcadores de neoplasias colorrectales de la Figura 6 (Panel A) y la Figura 13 (Panel B).

Los cebadores y las sondas pueden marcarse de forma detectable. Un marcado puede ser una fracción molecular o un compuesto que puede detectarse o dar lugar a una respuesta detectable, que puede unirse directa o indirectamente a un ácido nucleico. El etiquetado directo puede utilizar enlaces o interacciones para unir el marcado y la sonda, lo que incluye enlaces covalentes, interacciones no covalentes (enlaces de hidrógeno, interacciones hidrófobas e iónicas), o quelatos o complejos de coordinación. El etiquetado indirecto puede utilizar una fracción puente o un enlazador (por ejemplo, un anticuerpo, un oligómero u otro compuesto), etiquetado directa o indirectamente, que puede amplificar una señal. Las etiquetas incluyen cualquier fracción detectable, por ejemplo, radionucleido, ligando tal como biotina o avidina, enzima, sustrato enzimático, grupo reactivo, cromóforo (colorante, partícula o microesfera detectables), fluoróforo o compuesto luminiscente (etiqueta bioluminiscente, fosforescente o quimioluminiscente). Las etiquetas pueden ser detectables en un ensayo homogéneo en el que la sonda marcada unida en una mezcla presenta un cambio detectable en comparación con el de la sonda marcada no unida, por ejemplo, estabilidad o degradación diferencial, sin requerir la separación física de las formas unidas de las no unidas.

Las etiquetas detectables adecuadas pueden incluir moléculas que son detectables por sí mismas (por ejemplo, fracciones fluorescentes, etiquetas electroquímicas, quelatos metálicos, etc.), así como moléculas que pueden detectarse indirectamente mediante la producción de un producto de reacción detectable (por ejemplo, enzimas como la peroxidasa de rábano picante, la fosfatasa alcalina, etc.) o por una molécula de unión específica que puede ser detectable (por ejemplo, biotina, digoxigenina, maltosa, oligohistidina, 2,4-dinitrobenzenceno, fenilarsenato, ssADN, dsADN, etc.). Como se ha comentado anteriormente, el acoplamiento de uno o más motivos ligandos y/o ligandos al marcado detectable puede ser directo o indirecto. La detección puede ser in situ, in vivo, in vitro en una sección de tejido o en solución, etc.

En algunas realizaciones, los métodos incluyen el uso de sondas polinucleotídicas conjugadas con fosfatasa alcalina. Cuando se utiliza una sonda polinucleotídica conjugada con fosfatasa alcalina (AP), tras la adición secuencial de un sustrato apropiado, tal como el azul rápido o el rojo rápido, la AP descompone el sustrato para formar un precipitado que permite la detección in situ de la molécula de ARN diana específica. La fosfatasa alcalina puede utilizarse con

varios sustratos, por ejemplo, azul rápido, rojo rápido o 5-bromo-4-cloro-3-indolil-fosfato (BCIP). Véase, por ejemplo, como se describe en general en US 5,780,277 y US 7,033,758.

En algunas realizaciones, las sondas conjugadas con fluoróforo pueden ser sondas de etiquetas conjugadas con colorantes fluorescentes, o utilizar otros enfoques enzimáticos además de la fosfatasa alcalina para una ruta de detección cromogénica, como el uso de sondas conjugadas con peroxidasa de rábano picante con sustratos como 3,3'-diaminobencidina (DAB).

Los colorantes fluorescentes utilizados en las sondas de etiquetas conjugadas pueden dividirse típicamente en familias, tales como fluoresceína y sus derivados; rodamina y sus derivados; cianina y sus derivados; cumarina y sus derivados; Cascade Blue™ y sus derivados; Lucifer Yellow y sus derivados; BODIPY y sus derivados; y similares. Los fluoróforos ejemplares incluyen indocarbocianina (C3), indodicarbocianina (C5), Cy3, Cy3.5, Cy5, Cy5.5, Cy7, Texas Red, Pacific Blue, Oregon Green 488, Alexa Fluor®-355, Alexa Fluor 488, Alexa Fluor 532, Alexa Fluor 546, Alexa Fluor-555, Alexa Fluor 568, Alexa Fluor 594, Alexa Fluor 647, Alexa Fluor 660, Alexa Fluor 680, JOE, Lissamine, Rhodamine Green, BODIPY, isotiocianato de fluoresceína (FITC), carboxifluoresceína (FAM), ficoeritrina, rodamina, diclororodamina (dRhodamine™), carboxi tetrametilrodamina (TAMRA™), carboxi-X-rodamina (ROX™), LIZ™, VIC™, NED™, PET™, SYBR, PicoGreen, RiboGreen, y similares. Las descripciones de los fluoróforos y su uso pueden encontrarse, pero no se limitan a lugares, en, R Haugland, Handbook of Fluorescent Probes and Research Products, 9.ª ed. (2002), Molecular Probes, Eugene, Oreg.; M. Schena, Microarray Analysis (2003), John Wiley & Sons, Hoboken, N.J.; Synthetic Medicinal Chemistry 2003/2004 Catalog, Berry and Associates, Ann Arbor, Mich.; G. Hermanson, Bioconjugate Techniques, Academic Press (1996); y Glen Research 2002 Catalog, Sterling, Va. Los colorantes de infrarrojo cercano se incluyen expresamente en el significado de los términos fluoróforo y grupo reportero fluorescente.

Los niveles de los biomarcadores eucarióticos pueden analizarse en un arreglo de genes. El análisis de microarreglos puede realizarse en un arreglo de genes personalizado. Alternativamente, o además, el análisis de microarreglos puede llevarse a cabo utilizando sistemas disponibles en el mercado de acuerdo con las instrucciones y protocolos del fabricante. Algunos ejemplos de sistemas comerciales son la tecnología GENECHIP® de Affymetrix (Affymetrix, Santa Clara, CA), la tecnología de microarreglos de Agilent, el sistema de análisis NCOUNTER® (NanoString® Technologies) y la tecnología de microarreglos BeadArray (Illumina). Los ácidos nucleicos extraídos de una muestra de heces pueden hibridarse con las sondas del arreglo de genes. La hibridación sonda-destino puede detectarse mediante quimioluminiscencia para determinar la abundancia relativa de secuencias concretas.

Las sondas y conjuntos de sondas pueden configurarse como un arreglo de genes. Un arreglo de genes, también conocido como microarreglo o chip de genes, es un conjunto ordenado de ácidos nucleicos que permite el análisis paralelo de muestras biológicas complejas. Típicamente, un arreglo de genes incluye sondas que están unidas a un sustrato sólido, por ejemplo un microchip, un portaobjetos de vidrio o una perla. La fijación suele implicar un acoplamiento químico que da lugar a un enlace covalente entre el sustrato y la sonda. El número de sondas de un arreglo puede variar, pero cada sonda está fijada a una ubicación direccionable específica en el arreglo o microchip. En algunas realizaciones, las sondas pueden tener una longitud de aproximadamente 18 bases de nucleótidos, aproximadamente 20 bases de nucleótidos, aproximadamente 25 bases de nucleótidos, aproximadamente 30 bases de nucleótidos, aproximadamente 35 bases de nucleótidos o aproximadamente 40 bases de nucleótidos. En algunas realizaciones, el conjunto de sondas comprende sondas dirigidas a al menos 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 40, 45, 50, 60, 70, 80, 90, 100, 125, 150, o más, o todos, los genes biomarcadores de neoplasia colorrectal de la Figura 6 (Panel A) y la Figura 13 (Panel B). Los conjuntos de sondas pueden incorporarse en arreglos de alta densidad que comprenden 5.000, 10.000, 20.000, 50.000, 100.000, 200.000, 300.000, 400.000, 500.000, 600.000, 700.000, 800.000, 900.000, 1.000.000, 2.000.000, 3.000.000, 4.000.000, 5.000.000, 6.000.000, 7.000.000, 8.000.000 o más sondas diferentes.

Los métodos de síntesis de arreglos de genes pueden variar. Los métodos ejemplares incluyen la síntesis de las sondas seguida de su deposición en la superficie del arreglo por "depositado", síntesis in situ, utilizando, por ejemplo, fotolitografía, o electroquímica en arreglos de microelectrodos.

Los niveles de marcadores eucarióticos pueden evaluarse mediante diversos métodos. Los niveles de expresión pueden determinarse a nivel del ácido nucleico, por ejemplo, el ARN, o a nivel del polipéptido. La expresión de ARN puede abarcar la expresión de ARN total, ARNm, ARNt, ARNr, ARNnc, ARNsm, miARN y snoARN. La expresión a nivel de ARN puede medirse directa o indirectamente midiendo los niveles de ADNc correspondientes al ARN pertinente. Como alternativa, o además, también pueden analizarse los polipéptidos codificados por el ARN, los reguladores de ARN de los genes que codifican los factores de transcripción pertinentes, y los niveles de los polipéptidos del factor de transcripción. Los métodos para determinar la expresión génica a nivel del ARNm incluyen, por ejemplo, el análisis de microarreglos, el análisis en serie de la expresión génica (SAGE), la RT-PCR, la secuenciación, la hibridación basada en ensayos de cuantificación de códigos de barras digitales, la RT-PCR multiplex, la PCR de gota digital (ddPCR), los espectrofotómetros NanoDrop, la qRT-PCR, la qPCR, la espectroscopia UV, la secuenciación del ARN, la secuenciación de próxima generación, los ensayos de hibridación basados en lisados que utilizan la amplificación de señales de ADN ramificado, tales como el QuantiGene 2.0 Single Plex, y los métodos de análisis de ADN ramificado. Los ensayos digitales de cuantificación de códigos de barras pueden incluir el BeadArray (Illumina), los sistemas xMAP (Luminex), el nCounter (Nanostring), la genómica de alto rendimiento (HTG) molecular,

el BioMark (Fluidigm) o el microarreglo Wafergen. Los ensayos pueden incluir DASL (Illumina), RNA-Seq (Illumina), TruSeq (Illumina), SureSelect (Agilent), Bioanalyzer (Agilent) y TaqMan (ThermoFisher).

Los niveles de los biomarcadores eucarióticos también pueden analizarse mediante secuenciación del ADN. La secuenciación del ADN puede realizarse mediante métodos de secuenciación como la secuenciación dirigida, la secuenciación del genoma completo o la secuenciación del exoma. Los métodos de secuenciación pueden incluir: Secuenciación Sanger o secuenciación de alto rendimiento. La secuenciación de alto rendimiento puede incluir la secuenciación por síntesis, la pirosecuenciación, la secuenciación por ligadura, la secuenciación en tiempo real, la secuenciación por nanoporos y la secuenciación Sanger. En algunas realizaciones, el ARN aislado puede utilizarse para generar un ADNc correspondiente y el ADNc puede secuenciarse.

Los métodos de secuenciación pueden llevarse a cabo en formatos multiplex tales que múltiples ácidos nucleicos diana diferentes son manipulados simultáneamente. En algunas realizaciones, diferentes ácidos nucleicos diana pueden tratarse en un recipiente de reacción común o en una superficie de un sustrato particular, lo que permite una entrega conveniente de reactivos de secuenciación, la eliminación de reactivos sin reaccionar y la detección de eventos de incorporación de manera multiplex. En algunas realizaciones en las que están implicados ácidos nucleicos diana unidos a la superficie, los ácidos nucleicos diana pueden estar en un formato de arreglo. En un formato de arreglo, los ácidos nucleicos diana pueden estar típicamente acoplados a una superficie de forma espacialmente distinguible. Por ejemplo, los ácidos nucleicos diana pueden estar unidos por unión covalente directa, unión a una perla u otra partícula o asociados a una polimerasa u otra molécula que esté unida a la superficie. El arreglo puede incluir una sola copia de un ácido nucleico diana en cada sitio (también denominado característica) o pueden estar presentes en cada sitio o característica múltiples copias que tengan la misma secuencia. Las copias múltiples se producen mediante métodos de amplificación como la amplificación en puente, la PCR o la PCR en emulsión.

Se puede utilizar una etapa de normalización para controlar la recuperación de ácido nucleico y la variabilidad entre muestras. En algunas realizaciones, puede añadirse una cantidad definida de ácidos nucleicos de control exógenos ("añadidos como control") a los ácidos nucleicos eucariotas extraídos. El ácido nucleico de control exógeno puede ser un ácido nucleico que tenga una secuencia correspondiente a una o más secuencias eucariotas. Alternativamente, o además, el ácido nucleico de control exógeno puede tener una secuencia correspondiente a la secuencia encontrada en otra especie, por ejemplo una secuencia bacteriana como una secuencia de *Bacillus subtilis*. En algunas realizaciones, los métodos pueden incluir la determinación de los niveles de uno o más genes de mantenimiento. En algunas realizaciones, los métodos pueden incluir la normalización de los niveles de expresión de los biomarcadores a los niveles de los genes de referencia.

Los métodos pueden incluir la etapa de determinar si los niveles de expresión medidos de uno, dos o más biomarcadores en una muestra experimental son diferentes de los niveles de expresión medidos de uno, dos o más biomarcadores en un control. En otra realización, los métodos incluyen la etapa de determinar si la proporción de niveles de expresión de uno, dos o más biomarcadores en una muestra experimental es diferente de la proporción de niveles de expresión medidos de uno, dos o más biomarcadores en un control. Una diferencia en el nivel de expresión o una proporción de niveles de expresión puede ser un aumento o una disminución.

Las composiciones son generalmente y diversamente útiles para la detección, diagnóstico y tratamiento del cáncer colorrectal. Los métodos de detección pueden incluir la medición del nivel de expresión en una muestra de heces de dos o más biomarcadores de neoplasia colorrectal seleccionados entre los biomarcadores listados en cualquiera de las Figuras 6 (Panel A) o 13 (Panel B) y la comparación del nivel de expresión medido de los dos o más genes biomarcadores de neoplasia colorrectal en la muestra con el nivel de expresión medido de dos o más genes biomarcadores de neoplasia colorrectal en un control. Una diferencia en el nivel de expresión medido de dos o más genes biomarcadores de neoplasia colorrectal en la muestra de un paciente en relación con el nivel de expresión medido de los dos o más genes biomarcadores de neoplasia colorrectal en un control es una indicación de que el paciente tiene cáncer colorrectal. En algunas realizaciones, una diferencia en el nivel de expresión medido de dos o más genes biomarcadores de neoplasia colorrectal en la muestra de un paciente en relación con el nivel de expresión medido de los dos o más genes biomarcadores de neoplasia colorrectal en un control es una indicación de que el paciente tiene una lesión precancerosa y/o está en riesgo de cáncer colorrectal. Estos métodos pueden incluir además la etapa de identificar a un sujeto (por ejemplo, un paciente y, más específicamente, un paciente humano) que tiene una neoplasia colorrectal, por ejemplo, cáncer colorrectal o una lesión precancerosa, o que está en riesgo de desarrollar una neoplasia colorrectal.

Una neoplasia colorrectal puede incluir cualquier forma de cáncer colorrectal. Una neoplasia colorrectal también puede incluir un pólipo, por ejemplo una lesión precancerosa. El cáncer colorrectal suele comenzar como un crecimiento, denominado pólipo, en el revestimiento luminal del colon o el recto. Los pólipos colorrectales se dividen generalmente en dos categorías: pólipos adenomatosos, también llamados adenomas; y pólipos hiperplásicos e inflamatorios. Los pólipos adenomatosos pueden dar lugar a cáncer colorrectal. La forma más común de cáncer colorrectal, el adenocarcinoma, se origina en las células de las glándulas intestinales que recubren el interior del colon y/o el recto. Los adenocarcinomas pueden incluir adenocarcinomas tubulares, que son cánceres glandulares en un tallo pedunculado, y adenocarcinomas vellosos, que son cánceres glandulares que se sitúan planos en la superficie del colon. Otros cánceres colorrectales se distinguen por su tejido de origen. Entre ellos se incluyen los tumores del estroma gastrointestinal (GIST), que surgen de las células intersticiales de Cajal; los linfomas colorrectales primarios,

que surgen de células hematológicas; los leiomiomas, que son sarcomas que surgen del tejido conjuntivo o del músculo liso; los melanomas, que surgen de los melanocitos; carcinomas de células escamosas, que surgen del tejido epitelial escamoso estratificado y se limitan al recto; y carcinomas mucinosos, que son cánceres epiteliales generalmente asociados a un mal pronóstico.

- 5 Los síntomas del cáncer colorrectal pueden incluir, pero no se limitan a, un cambio en los hábitos intestinales, incluyendo diarrea o estreñimiento o un cambio en la consistencia de las heces que dure más de cuatro semanas, sangrado rectal o sangre en las heces, molestias abdominales persistentes como calambres, gases o dolor, sensación de que el intestino no se vacía completamente, debilidad o fatiga y pérdida de peso inexplicable. Los pacientes con sospecha de cáncer colorrectal pueden someterse a análisis de sangre periférica, como un hemograma completo (CBC), una prueba de sangre oculta en heces (FOBT), un análisis de la función hepática, una prueba inmunoquímica fecal (FIT) y/u otros análisis de ciertos marcadores tumorales, por ejemplo el antígeno carcinoembrionario (CEA) y el CA19-9. El cáncer colorrectal suele diagnosticarse mediante colonoscopia. Durante la colonoscopia, se extirpan los pólipos detectados, se les practica una biopsia y se analizan para determinar si contienen células cancerosas colorrectales o células que han sufrido un cambio precanceroso. Cada uno de los cánceres específicos listados anteriormente puede tener un aspecto diferente cuando se observa a través de un endoscopio. Los adenomas vellosos, los melanomas y los carcinomas de células escamosas suelen ser planos o sésiles, mientras que los adenomas tubulares, los linfomas, los leiomiomas y los tumores GIST suelen ser pediculados. Sin embargo, los adenomas planos y sésiles pueden pasar desapercibidos para los gastroenterólogos durante las colonoscopias. Las muestras de biopsia pueden someterse a análisis adicionales basados en cambios genéticos de genes concretos o en la inestabilidad de microsátélites.

Otros métodos de diagnóstico pueden incluir, sigmoidoscopia, pruebas de imagen, por ejemplo, escáneres de tomografía computarizada (CT o CAT); ultrasonido, por ejemplo abdominal, endorrectal o ultrasonido intraoperatorio, escáneres para imágenes de resonancia magnética (MRI), por ejemplo MRI endorrectal. Pueden realizarse otras pruebas, como angiografías y radiografías de tórax, para determinar si un cáncer colorrectal ha hecho metástasis.

- 25 Se han desarrollado diversos métodos para la estadificación del cáncer colorrectal. El sistema más comúnmente utilizado, el sistema TNM, se basa en tres factores: 1) la distancia a la que el tumor primario (T) ha crecido en la pared del intestino y zonas cercanas; 2) si el tumor se ha extendido a los ganglios linfáticos regionales cercanos (N); 3) si el cáncer ha hecho metástasis a otros órganos (M). Otros métodos de estadificación son la estadificación de Dukes y la clasificación de Astler-Coller.
- 30 El sistema TNM proporciona una clasificación en cuatro estadios del cáncer colorrectal. En el cáncer colorrectal en estadio 1 (T1), el tumor ha crecido en las capas de la pared del colon, pero no se ha extendido fuera de la pared del colon ni a los ganglios linfáticos. Si el cáncer forma parte de un pólipo de adenoma tubular, se realiza una escisión simple y el paciente puede seguir sometiéndose a pruebas rutinarias para detectar el desarrollo futuro de cáncer. Si el cáncer es de alto grado o forma parte de un pólipo plano o sésil, puede ser necesaria más cirugía y se tomarán márgenes más amplios; esto puede incluir una colectomía parcial en la que se reseca una sección del colon. En el cáncer colorrectal en estadio 2 (T2), el tumor ha crecido hacia la pared del colon y, potencialmente, hacia los tejidos cercanos, pero no se ha extendido a los ganglios linfáticos próximos. Generalmente se realiza la extirpación quirúrgica del tumor y una colectomía parcial. Puede administrarse terapia complementaria, por ejemplo, quimioterapia con agentes como 5-fluorouracilo, leucovorina o capecitabina. Es poco probable que estos tumores reaparezcan, pero en general es necesario aumentar la vigilancia del paciente. En el cáncer colorrectal en estadio 3 (T3), el tumor se ha extendido a los ganglios linfáticos cercanos, pero no a otras partes del cuerpo. Será necesaria una intervención quirúrgica para extirpar la sección del colon y todos los ganglios linfáticos afectados. Normalmente se recomienda la quimioterapia, con agentes como 5-fluorouracilo, leucovorina, oxaliplatino o capecitabina combinada con oxaliplatino. También puede utilizarse radioterapia, dependiendo de la edad del paciente y de la agresividad del tumor. En el cáncer colorrectal en estadio 4 (T4), el tumor se ha extendido desde el colon a órganos distantes a través de la sangre. El cáncer colorrectal suele metastatizar en el hígado, los pulmones y/o el peritoneo. Es poco probable que la cirugía cure estos cánceres y, por lo general, se necesita quimioterapia o radioterapia para mejorar las tasas de supervivencia.

Los métodos son generalmente útiles para el diagnóstico y tratamiento del cáncer colorrectal. El nivel de dos o más genes biomarcadores de neoplasia colorrectal se mide en una muestra biológica, es decir, una muestra de un sujeto.

- 50 El sujeto puede ser un paciente con uno o más de los síntomas descritos anteriormente que indicarían que el paciente está en riesgo de cáncer colorrectal. El sujeto también puede ser un paciente asintomático, pero con riesgo de cáncer colorrectal en función de su edad (por ejemplo, por encima de los 50 años), antecedentes familiares, obesidad, dieta, consumo de alcohol, consumo de tabaco, diagnóstico previo de pólipos colorrectales, raza y origen étnico, la enfermedad inflamatoria intestinal y los síndromes genéticos, como la poliposis adenomatosa familiar, el síndrome de Gardner, el síndrome de Lynch, el síndrome de Turcot, el síndrome de Peutz-Jeghers y la poliposis asociada a MUTYH, asociados a un mayor riesgo de cáncer colorrectal. Los métodos aquí divulgados también son útiles para monitorizar a un paciente que ha sido diagnosticado y tratado previamente de cáncer colorrectal con el fin de monitorizar la remisión y detectar la recurrencia del cáncer.

- 60 El estado de enfermedad de un sujeto, es decir, un paciente humano o animal no humano, se determina mediante una evaluación patológica. Por ejemplo, en un tipo de enfermedad, como el cáncer colorrectal, la extensión de la enfermedad se clasifica como estadio 1 (T1), estadio 2 (T2), estadio 3 (T3) y estadio 4 (T4). El cáncer colorrectal puede

ser un adenocarcinoma tubular, un adenocarcinoma vellosa, un tumor del estroma gastrointestinal, un linfoma colorrectal primario, un leiomioma, un melanoma, un carcinoma de células escamosas o un carcinoma mucinoso. En otro tipo de enfermedades, tal como la enfermedad inflamatoria intestinal, el estado de la enfermedad viene determinado por la localización de la enfermedad a lo largo del tracto intestinal y por características histológicas como granulomas, infiltrados leucocitarios y/o abscesos de criptas. Otros métodos para determinar el estado de la enfermedad, como la determinación del médico, los síntomas físicos, una prueba de sangre oculta en heces, una prueba inmunológica fecal, una sigmoidoscopia, FIT-DNA, una colonografía por TAC o una colonoscopia, también pueden utilizarse junto con los métodos aquí divulgados.

La enfermedad intestinal puede incluir cáncer intestinal, cáncer colorrectal, pólipos adenomatosos indicativos de cambio precanceroso, síndrome del intestino irritable, colitis ulcerosa, enfermedad de Crohn u otra enfermedad intestinal. Un método para determinar si un sujeto está en riesgo de padecer una enfermedad intestinal puede determinarse utilizando la invención para detectar a) una secuencia de ácido desoxirribonucleico (ADN), b) una secuencia de ácido ribonucleico (ARN), c) una secuencia de aminoácidos predicha, que comprende la columna vertebral de la proteína, d) niveles de expresión de biomarcadores de ácido ribonucleico, e) predicción en la variación de una secuencia en aminoácido o f) cualquier combinación de las anteriores, donde una diferencia entre el control y la muestra experimental puede indicar que el sujeto está en riesgo de padecer una enfermedad intestinal.

Los métodos de la invención y las composiciones aquí descritas también son útiles para seleccionar un plan clínico para un sujeto que padece una enfermedad intestinal. Mediante este método, el plan clínico puede incluir la administración de otros procedimientos diagnósticos. En algunas realizaciones, el plan clínico puede incluir un método de tratamiento.

El nivel de dos o más genes biomarcadores de neoplasia colorrectal seleccionados de la Figura 6 (Panel A) o la Figura 13 (Panel B) puede analizarse en un sujeto con riesgo de padecer cáncer colorrectal o que lo padece. En algunas realizaciones, el nivel de uno, dos o más genes biomarcadores de neoplasia colorrectal seleccionados de la Figura 6 (Panel A) o la Figura 13 (Panel B) puede analizarse en un sujeto con riesgo de padecer cáncer colorrectal o que lo padece. Los genes biomarcadores de neoplasias colorrectales listados en la Figura 6 (Panel A) o en la Figura 13 (Panel B) pueden formar un panel. En algunas realizaciones, los dos o más biomarcadores pueden incluir combinaciones de 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 120, 140, 160, 180 o más de los marcadores de la Figura 6 (Panel A) o la Figura 13 (Panel B). En algunas realizaciones, los marcadores pueden estar contenidos dentro de grupos de transcripción expresados diferencialmente y/o vías comunes asociadas con el cáncer colorrectal. Algunas vías ejemplares incluyen la inestabilidad de microsatélites (MSI), la inestabilidad cromosómica (CIN) y el fenotipo metilador de islas CpG (CIMP). En algunas realizaciones, las vías pueden ser vías de componentes celulares, de respuesta celular al estrés, de estrés y de unión de ARN.

Los algoritmos para determinar el diagnóstico, el estado o la respuesta al tratamiento, por ejemplo, pueden determinarse para condiciones clínicas particulares. Los algoritmos utilizados en los métodos aquí proporcionados pueden ser funciones matemáticas que incorporan múltiples parámetros que pueden cuantificarse utilizando, sin limitación, dispositivos médicos, puntuaciones de evaluación clínica o pruebas biológicas/químicas/físicas de muestras biológicas. Cada función matemática puede ser una expresión ajustada al peso de los niveles (por ejemplo, niveles medidos) de parámetros determinados como relevantes para una condición clínica seleccionada. Debido a las técnicas implicadas en la ponderación y evaluación de paneles de marcadores múltiples, pueden utilizarse ordenadores con una potencia de cálculo razonable para analizar los datos.

Un método de diagnóstico puede incluir la obtención de una muestra de heces de un paciente con riesgo o sospecha de cáncer colorrectal; determinar la expresión de dos o más genes biomarcadores de neoplasia colorrectal seleccionados de la Figura 6 (Panel A) o la Figura 13 (Panel B) y proporcionar un valor de prueba mediante los algoritmos de aprendizaje automático que incorporan una pluralidad de genes biomarcadores de neoplasia colorrectal seleccionados de cualquiera de los paneles de genes biomarcadores de neoplasia colorrectal con un coeficiente predefinido. Los algoritmos de ejemplo de aprendizaje automático son Support Vector Machine, Gradient Boosting, Adaptive Boosting, Random Forest, Naive Bayes, Decision Tree y k-Nearest Neighbors. Un cambio significativo en la expresión de una pluralidad de genes biomarcadores de neoplasia colorrectal en relación con el control, por ejemplo, una población de individuos sanos, indica una mayor probabilidad de que el paciente tenga cáncer colorrectal y/o una lesión precancerosa. En algunas realizaciones, los niveles de expresión medidos en una muestra se utilizan para derivar o calcular una probabilidad o una puntuación de confianza. Este valor puede derivarse de los niveles de expresión. Alternativamente, o además, el valor puede derivarse de una combinación del valor de expresión con otros factores, por ejemplo, el historial médico, la edad y los antecedentes genéticos del paciente. En algunas realizaciones, el método puede comprender además la etapa de comunicar el valor de la prueba al paciente.

Los dispositivos y sistemas informáticos estándar pueden ser utilizados e implementados, por ejemplo, convenientemente programados, para realizar los métodos aquí descritos, por ejemplo, para realizar los cálculos necesarios para determinar los valores aquí descritos. Los dispositivos informáticos incluyen diversas formas de ordenadores digitales, como portátiles, ordenadores de sobremesa, dispositivos móviles, estaciones de trabajo, asistentes digitales personales, servidores, servidores blade, mainframes y otros ordenadores apropiados. En algunas

realizaciones, el dispositivo informático es un dispositivo móvil, tal como un asistente digital personal, un teléfono móvil, un smartphone, una tableta u otro dispositivo informático similar.

En algunas realizaciones, se puede utilizar un ordenador para comunicar información, por ejemplo, a un profesional sanitario. Se puede comunicar información a un profesional poniendo dicha información a su disposición por vía electrónica (por ejemplo, de forma segura). Por ejemplo, la información puede almacenarse en una base de datos informática para que un profesional sanitario pueda acceder a ella. Además, la información puede comunicarse a un hospital, clínica o centro de investigación que actúe como agente del profesional. La información transferida a través de redes abiertas (por ejemplo, Internet o correo electrónico) puede cifrarse. Los datos y análisis de la expresión génica del paciente pueden almacenarse en la nube con cifrado. El método AES de 256 bits con protección antimanipulación puede utilizarse para el cifrado de discos; el protocolo SSL, preferiblemente, puede garantizar la protección en el tránsito de datos, y la técnica de gestión de claves SHA2-HMAC puede permitir el acceso autenticado a los datos. También pueden utilizarse otros medios de almacenamiento seguro de datos.

Los resultados de dicho análisis pueden servir de base para el seguimiento y el tratamiento por parte del médico. Si el nivel de expresión de dos o más genes biomarcadores de neoplasias colorrectales seleccionados de la Figura 6 (Panel A) o de la Figura 13 (Panel B) no es significativamente diferente del nivel de expresión de los mismos dos o más biomarcadores de neoplasias colorrectales en un control, el médico puede determinar que el paciente no está actualmente en riesgo de cáncer colorrectal o de una lesión precancerosa. Se puede animar a estos pacientes a que vuelvan en el futuro para someterse a un nuevo cribado. Los métodos aquí descritos pueden utilizarse para monitorizar cualquier cambio en los niveles de los marcadores de neoplasia colorrectal a lo largo del tiempo. Un sujeto puede ser sometido a seguimiento durante cualquier periodo de tiempo tras el cribado y/o diagnóstico inicial. Por ejemplo, un sujeto puede ser monitorizado durante al menos 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 45, 50, 55, o 60 meses o más o durante al menos 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 o más años.

Los métodos de la presente invención y las composiciones aquí descritas son útiles para seleccionar un plan clínico para un sujeto en riesgo de padecer o que padece una neoplasia colorrectal. El plan clínico puede incluir la administración de otros procedimientos diagnósticos, por ejemplo, una prueba de sangre oculta en heces, una prueba inmunoquímica fecal o una colonoscopia para extirpar pólipos o lesiones precancerosas. En algunas realizaciones, el plan clínico puede incluir un método de tratamiento. En algunas realizaciones, los métodos incluyen métodos de selección de un tratamiento para un sujeto que tiene cáncer colorrectal. Si el nivel de expresión de dos o más genes biomarcadores de neoplasia colorrectal seleccionados de la Figura 6 (Panel A) o la Figura 13 (Panel B) es significativamente diferente del nivel de expresión de los mismos dos o más genes biomarcadores de neoplasia colorrectal en un control, el paciente puede tener cáncer colorrectal. En estos casos, puede recomendarse un cribado adicional, por ejemplo, una mayor frecuencia de cribado utilizando los métodos aquí divulgados, así como una prueba de sangre oculta en el feto, una prueba inmunoquímica fecal y/o una colonoscopia. En algunos casos, se puede recomendar un tratamiento que incluya, por ejemplo, una colonoscopia con extirpación de pólipos, quimioterapia o cirugía, como la resección intestinal. Así, los métodos pueden utilizarse para determinar el nivel de expresión de dos o más genes biomarcadores de neoplasias colorrectales y, a continuación, determinar un curso de tratamiento. Un sujeto, es decir, un paciente, recibe un tratamiento eficaz siempre que se obtiene un resultado clínicamente beneficioso. Esto puede significar, por ejemplo, una resolución completa de los síntomas de una enfermedad, una disminución de la gravedad de los síntomas de la enfermedad o una ralentización de la progresión de la enfermedad. Estos métodos pueden incluir además las etapas de a) identificar a un sujeto (por ejemplo, un paciente y, más específicamente, un paciente humano) que tiene cáncer colorrectal; y b) proporcionar al sujeto un tratamiento contra el cáncer, por ejemplo, un agente terapéutico, cirugía o radioterapia. Se considera una cantidad terapéuticamente eficaz la cantidad de un agente terapéutico administrada al sujeto que produce una resolución completa de los síntomas de una enfermedad, una disminución de la gravedad de los síntomas de la enfermedad o una ralentización de la progresión de la enfermedad. Los métodos actuales también pueden incluir una etapa de monitorización para ayudar a optimizar la dosificación y la programación, así como predecir el resultado. La monitorización también puede utilizarse para detectar la aparición de resistencia al fármaco, distinguir rápidamente a los pacientes que responden de los que no o evaluar la recurrencia de un cáncer. Cuando hay signos de resistencia o falta de respuesta, el médico puede elegir un agente alternativo o complementario antes de que el tumor desarrolle mecanismos de escape adicionales.

Los métodos aquí divulgados también pueden utilizarse en combinación con métodos convencionales para el diagnóstico y tratamiento del cáncer colorrectal. Por lo tanto, los métodos de diagnóstico pueden utilizarse junto con los métodos de diagnóstico estándar para el cáncer colorrectal. Por ejemplo, los métodos pueden utilizarse en combinación con una prueba de sangre oculta en heces, una prueba inmunoquímica fecal o una colonoscopia. Los métodos también pueden utilizarse con otros marcadores de cáncer colorrectal, por ejemplo, KRAS, NRAS, BRAF, CEA, CA 19-9, p53, MSL, DCC y MMR.

Los métodos de diagnóstico también pueden utilizarse en combinación con tratamientos contra el cáncer colorrectal. Los métodos de tratamiento del cáncer colorrectal se dividen en varias categorías generales: cirugía, quimioterapia, radioterapia, terapia dirigida e inmunoterapia. La cirugía puede incluir colectomía, colostomía junto con hepatectomía parcial o proctectomía. La quimioterapia puede ser quimioterapia sistémica o quimioterapia regional, en la que los agentes quimioterapéuticos se colocan en proximidad directa a un órgano afectado. Los agentes quimioterapéuticos ejemplares pueden incluir 5-fluorouracilo, oxaliplatino o derivados del mismo, irinotecán o un derivado del mismo,

leucovorina o capecitabina, mitomicina C, cisplatino y doxorrubicina. La radioterapia puede ser radioterapia externa, en la que se utiliza una máquina para dirigir la radiación hacia el cáncer, o radioterapia interna, en la que se coloca una sustancia radiactiva directamente en el cáncer colorrectal o cerca de él. Los agentes diana pueden incluir agentes antiangiogénicos como bevacizumab) o anticuerpo monoclonal inhibidor del EGFR (cetuximab, panitumumab), ramucirumab (anti-VEGFR2), aflibercept, regorafenib, tripfluridina-tipiracilo o una combinación de los mismos. Los agentes diana también pueden combinarse con agentes quimioterapéuticos estándar. La inmunoterapia puede incluir la administración de anticuerpos específicos, por ejemplo anticuerpos anti-PD-1, anticuerpos anti-PD-L1, y anticuerpos temporales-CTLA-4, anticuerpos anti-CD 27; vacunas contra el cáncer, terapia celular adoptiva, terapias con virus oncolíticos, inmunoterapias adyuvantes y terapias basadas en citocinas. Otros métodos de tratamiento son el trasplante de células madre, la hipertermia, la terapia fotodinámica, la donación y transfusión de hemoderivados o el tratamiento con láser.

Podemos utilizar los términos “aumentado”, “incremento” o “regulado al alza” para referirnos en general a un aumento del nivel de un biomarcador en una cantidad estadísticamente significativa. En algunas realizaciones, un aumento puede ser un aumento de al menos el 10 % en comparación con un control, por ejemplo un aumento al menos aproximadamente del 20 %, o al menos aproximadamente del 30 %, o al menos aproximadamente del 40 %, o al menos aproximadamente del 50 %, o al menos aproximadamente del 60 %, o al menos aproximadamente del 70 %, o al menos aproximadamente del 80 %, o al menos aproximadamente del 90 % o hasta e incluyendo un aumento del 100 % o cualquier aumento entre el 10-100 % en comparación con un control, o al menos aproximadamente 0.5 veces, o al menos aproximadamente 1.0 veces, o al menos aproximadamente 1.2 veces, o al menos aproximadamente un aumento de 1.5 veces, o al menos aproximadamente un aumento de 2 veces, o al menos aproximadamente un aumento de 3 veces, o al menos aproximadamente un aumento de 4 veces, o al menos aproximadamente un aumento de 5 veces, o al menos aproximadamente un aumento de 10 veces, o cualquier aumento entre 1.0 veces y 10 veces o más en comparación con un control

Podemos utilizar los términos “disminución”, “disminuido”, “reducido”, “reducción” o “regulado a la baja” para referirnos a una disminución del nivel de un biomarcador eucariota en una cantidad estadísticamente significativa. En algunas realizaciones, una disminución puede ser una disminución de al menos 10% en comparación con un control, por ejemplo una disminución de al menos aproximadamente 20 %, o al menos aproximadamente 30 %, o al menos aproximadamente 40 %, o al menos aproximadamente 50 %, o al menos aproximadamente 60 %, o al menos aproximadamente 70 %, o al menos aproximadamente 80 %, o al menos aproximadamente 90 % o hasta e incluyendo una disminución del 100 % (es decir en comparación con un control), o cualquier disminución entre el 10-100 % en comparación con un control, o al menos aproximadamente unas 0.5 veces, o al menos aproximadamente unas 1.0 veces, o al menos aproximadamente unas 1.2 veces, o al menos aproximadamente unas 1.5 veces, o al menos aproximadamente unas 2 veces, o al menos aproximadamente unas 3 veces, o al menos aproximadamente unas 4 veces, o al menos aproximadamente unas 5 veces, o al menos aproximadamente unas 10 veces, o cualquier disminución entre 1.0 veces y 10 veces o mayor en comparación con un control.

La significancia estadística de un aumento de un biomarcador eucariota o de una disminución de un biomarcador eucariota puede expresarse como un valor p. Dependiendo del biomarcador eucariota específico, el valor p puede ser inferior a 0.05, inferior a 0.01, inferior a 0.005, inferior a 0.002, inferior a 0.001 o inferior a 0.0005.

Un control puede ser una muestra biológica obtenida de un paciente o de un grupo de pacientes. En algunas realizaciones, el control puede ser un valor de referencia. Un control puede obtenerse de un individuo, o de una población de individuos, que hayan sido diagnosticados como sanos. Los individuos sanos pueden incluir, por ejemplo, individuos que han dado negativo en una prueba parasitaria fecal, una prueba de bacterias fecales o una endoscopia en el último año. Un control puede obtenerse de un individuo, o de una población de individuos, que hayan sido diagnosticados como enfermos. Los individuos enfermos pueden incluir, por ejemplo, individuos que hayan dado positivo en una prueba parasitaria fecal, una prueba bacteriana fecal o una endoscopia en el último año. Un control puede obtenerse de un individuo, o de una población de individuos, a los que previamente se les había diagnosticado la enfermedad, pero que actualmente están en remisión o no padecen la enfermedad. Se puede obtener un control de un individuo en uno, dos o más momentos. Por ejemplo, un control puede ser una muestra biológica obtenida de un sujeto en un momento anterior. Un control puede ser un valor de referencia estándar para un biomarcador concreto. Puede obtenerse un valor de referencia estándar a partir de la evaluación de individuos de edad, sexo, género, tamaño corporal, raza, origen étnico o estado de salud general similar.

Una muestra experimental puede ser una muestra biológica obtenida de un sujeto. Una muestra experimental puede obtenerse de un sujeto con un estado de salud conocido o desconocido. En algunas realizaciones, el estado de salud de un sujeto puede determinarse, por ejemplo, mediante el análisis de una muestra experimental, biopsia, examen físico, hallazgos de laboratorio, inspección visual o análisis genético. El estado de salud de un sujeto que puede determinarse mediante una muestra experimental puede ser de enfermedad, de riesgo de enfermedad o de salud.

Artículos de fabricación

Los kits para detectar y cuantificar biomarcadores de neoplasias colorrectales seleccionados en una muestra biológica, por ejemplo, una muestra de heces, se refieren a productos envasados (por ejemplo, recipientes estériles que contienen una o más de las composiciones y se envasan para su almacenamiento, envío o venta en concentraciones

concentradas o listas para su uso) y kits. Un producto puede incluir un recipiente (por ejemplo, un vial, jarra, botella, bolsa, microplaca, microchip o perlas) que contenga una o más composiciones. Además, un artículo de fabricación puede incluir, por ejemplo, materiales de envasado, instrucciones de uso, jeringas, dispositivos de administración, tampones u otros reactivos de control.

5 El kit puede incluir un compuesto o agente capaz de detectar ARN correspondiente a dos o más de los genes biomarcadores de neoplasia colorrectal seleccionados de la Figura 6 (Panel A) o la Figura 13 (Panel B) en una muestra biológica; y un estándar; y opcionalmente uno o más reactivos necesarios para realizar la detección, cuantificación o amplificación. El kit puede incluir un compuesto o agente capaz de detectar ARN correspondiente a dos o más de un
10 marcador de células B, un marcador de células T, o una inmunoglobulina, en una muestra biológica; y un estándar; y opcionalmente uno o más reactivos necesarios para realizar la detección, cuantificación o amplificación. Los compuestos, agentes y/o reactivos pueden empacarse en un recipiente adecuado. El kit puede comprender además instrucciones de uso para detectar y cuantificar el ácido nucleico. Por ejemplo, el kit puede incluir: (1) una sonda, por ejemplo, un oligonucleótido, por ejemplo un oligonucleótido marcado detectablemente, que se hibrida con una
15 secuencia de ácido nucleico correspondiente a dos o más de los genes biomarcadores colorrectales seleccionados de la Figura 6 (Panel A) o la Figura 13 (Panel B) o un marcador de células B, un marcador de células T, o una inmunoglobulina o (2) un par de cebadores útiles para amplificar una molécula de ácido nucleico correspondiente a dos o más de los genes biomarcadores colorrectales seleccionados de entre los de la Figura 6 (Panel A) o la Figura 13 (Panel B) o un marcador de células B, un marcador de células T o una inmunoglobulina. El kit puede incluir además
20 sondas y cebadores útiles para amplificar uno o más genes de referencia. El kit también puede incluir un agente amortiguador, un conservante y/o un agente estabilizador de ácidos nucleicos o proteínas. El kit también puede incluir los componentes necesarios para detectar el agente detectable (por ejemplo, una enzima o un sustrato). El kit también puede contener un control o una serie de controles que pueden analizarse y compararse con la muestra de ensayo contenida. Cada componente del kit puede estar incluido en un empaque individual y todos los diversos envases pueden estar dentro de un único paquete, junto con instrucciones para interpretar los resultados de los ensayos
25 realizados utilizando el kit. Los kits pueden incluir cebadores o sondas de oligonucleótido específicas para uno o más marcadores de control. Los kits incluyen reactivos específicos para la cuantificación de dos o más de los biomarcadores colorrectales seleccionados de la Figura 6 (Panel A) o la Figura 13 (Panel B) o un marcador de células B, un marcador de células T o una inmunoglobulina.

30 El kit puede incluir reactivos específicos para la separación de células humanas de células bacterianas y otros componentes de las heces y la extracción de ARNm humano de la muestra de heces de un paciente. Así, el kit puede incluir tampones, perlas de emulsión, perlas de sílice, reactivos de estabilización y diversos filtros y recipientes para la centrifugación. El kit también puede incluir instrucciones de manipulación de las heces para minimizar la contaminación de las muestras y garantizar la estabilidad del ARNm humano en la muestra de heces. El kit también puede incluir ítems para garantizar la conservación de la muestra, por ejemplo, refrigerantes o paquetes térmicos. El
35 kit puede incluir un dispositivo de recogida de heces.

El producto también puede incluir una leyenda (por ejemplo, un marcado o inserto impreso u otro medio que describa el uso del producto (por ejemplo, una cinta de audio o vídeo o un medio legible por ordenador)). La leyenda puede estar asociada al envase (por ejemplo, adherida al envase) y puede describir la forma en que pueden utilizarse los reactivos. Los reactivos pueden estar listos para su uso (por ejemplo, presentes en unidades apropiadas), y pueden
40 incluir uno o más adyuvantes adicionales, portadores u otros diluyentes. Alternativamente, los reactivos pueden suministrarse en forma concentrada con un diluyente e instrucciones para la dilución.

Ejemplos

Ejemplo 1: Obtención de muestras de heces humanas

45 Recogida de heces humanas: Se pidió a los pacientes que defecaran en un cubo que cabía sobre el asiento de un retrete y las muestras resultantes se almacenaron en un congelador hasta que se transportaron a la Universidad Médica Nacional de Kharkiv (Kharkiv, Ucrania). Las heces se alicuotaron en tubos cónicos de 50 ml y se almacenaron a -80°C. Las muestras se enviaron desde la Universidad Médica Nacional de Kharkiv en hielo seco a Capital Biosciences (Gaithersburg, MD) y se transfirieron inmediatamente a un congelador a -80 °C. Desde allí, las muestras se enviaron en hielo seco a BioGenerator Labs (Saint Louis, MO), donde se almacenaron en un congelador a -80 °C
50 hasta su extracción.

Tipos de muestras humanas: Se obtuvieron muestras de heces de 195 pacientes con cáncer colorrectal (estadios I-IV), 126 pacientes con adenomas precancerosos, 125 pacientes con resultados negativos en una colonoscopia y 8
55 pacientes con pólipos benignos, lo que dio lugar a 454 muestras agregadas. Los individuos sanos eran pacientes sin antecedentes de cáncer colorrectal, enfermedad inflamatoria intestinal, celiaquía, síndrome del intestino irritable, diarrea en los últimos 20 días o cualquier otra enfermedad gastrointestinal. Los individuos enfermos eran pacientes diagnosticados de cáncer colorrectal y pólipos precancerosos. A los pacientes con cáncer colorrectal se les había diagnosticado cáncer colorrectal en estadio I-estadio IV mediante colonoscopia y posterior biopsia en el último mes y aún no habían recibido ningún tratamiento posterior a la biopsia, que puede incluir quimioterapia, radioterapia y/o
60 cirugía. Los pacientes con pólipos proporcionaron una muestra de heces antes de someterse a una colonoscopia en la que el médico detectó un pólipo que se consideró precanceroso mediante una posterior biopsia y evaluación

histológica. Los individuos sanos se emparejaron con los pacientes con pólipos y cáncer en función del sexo y los tramos de edad (50-60 años, 60-70 años, 70-80 años y 80-90 años). Los pacientes utilizados para este estudio recibieron el consentimiento de Capital Biosciences. La Junta de Revisión Interna de Schulman se encargó de la supervisión ética de este estudio.

5 Ejemplo 2: Extracción de ácido nucleico humano

Extracción total de ácido nucleico: Cada muestra de heces se colocó en un tubo cónico de 50 ml. Se añadieron aproximadamente 1.000-25.000 mg de heces a cada tubo. Se añadieron otros 20-40 ml de solución a cada tubo. Esta solución contenía una mezcla de solución salina equilibrada de Hanks (HBSS) (Sigma-Aldrich) con 0.05 % de Tween-20 (Sigma-Aldrich) y 0.0002 % de inhibidor de RNase (Sigma-Aldrich). Las heces se suspendieron en la solución y se rotaron a aproximadamente 0-10 °C durante 0-10 minutos. La solución se centrifugó a 1000 rpm a 4 °C durante 10 minutos y se desechó el sobrenadante. Se añadieron aproximadamente 4-10 ml de tampón de lisis EasyMag® (bioMerieux) al gránulo y éste se volvió a suspender en la solución. La solución se centrifugó a 2500-3500 rpm a 20-25 °C durante 10-15 minutos. Durante la centrifugación diferencial, la solución se separó en tres capas. La capa inferior incluía restos celulares sólidos, la capa intermedia era una capa hidrófila enriquecida en ácido nucleico humano y la capa superior era una capa lipídica hidrófoba. Las dos capas superiores se transfirieron a un nuevo tubo cónico de 15 ml y la solución se centrifugó de nuevo a 2500 rpm a 20-25 °C durante 10 minutos. El resultado de esta etapa de centrifugación fue la separación en tres capas: la capa inferior eran restos celulares sólidos, la capa intermedia era una capa hidrófila enriquecida en ácido nucleico humano y la capa superior era una capa lipídica hidrófoba. Para filtrar los residuos grandes de la solución, se colocó una punta de pipeta de 20 µl en una punta de pipeta de 1 ml y se pipetearon 2 ml de la capa hidrófila del tubo de 15 ml y se transfirieron a un cartucho desechable EasyMag® (bioMerieux). Además, se añadieron 60 µl de sílice magnética EasyMag® (bioMerieux) al cartucho. Las perlas se mezclaron en la solución durante 0.5-1 minutos utilizando una pipeta. Los ácidos nucleicos, que estaban unidos a las perlas, se eluyeron en una solución tampón utilizando el Protocolo A Específico de acuerdo con las instrucciones del fabricante. El volumen de los ácidos nucleicos eluidos fue de 70 µl. Esta solución de ácido nucleico se pipeteó en un tubo de 1.5 ml y se colocó en hielo. Los mismos cartuchos desechables EasyMag® (bioMerieux) que se utilizaron en la etapa anterior se volvieron a cargar con 2 ml adicionales de la capa hidrófila de la misma solución en el tubo de 15 ml utilizado anteriormente utilizando la misma técnica para filtrar los residuos grandes. Se añadieron 20 µl adicionales de sílice magnética EasyMag® (bioMerieux) al cartucho. Las perlas se mezclaron en la solución durante 0.5-1 minutos utilizando una pipeta. Como se ha descrito anteriormente, los ácidos nucleicos, que estaban unidos a las perlas, se eluyeron en una solución tampón utilizando el Protocolo A Específico de acuerdo con las instrucciones del fabricante. El volumen de los ácidos nucleicos eluidos fue de 70 µl. Esta solución de ácido nucleico se pipeteó en el tubo original de 1.5 ml que ya contenía el primer eluido de 70 µl y la solución combinada se colocó en hielo.

Tratamiento con DNasa: La solución de 140 µl se trató con Baseline-Zero- DNasa (Epicenter) a 35-40 °C durante 20-40 minutos. Se añadió una alícuota de 1-2 ml de tampón de lisis EasyMag® a la solución tratada con DNasa y se transfirió la muestra a un nuevo cartucho EasyMag® desechable. Se añadió toda la solución al nuevo cartucho junto con 60 µl de sílice magnética EasyMag®. Los ácidos nucleicos, que estaban unidos a las perlas, se eluyeron en una solución tampón utilizando el protocolo EasyMag® Generic de acuerdo con las instrucciones del fabricante. El volumen de los ácidos nucleicos eluidos fue de 25 µl. Esta solución de ácido nucleico se pipeteó en un tubo de 1.5 ml y se almacenó a 0-6 °C.

40 Ejemplo 3: Medición de los niveles de ácido nucleico humano en muestras de heces humanas

Resultados de la extracción: Utilizando la muestra extraída anteriormente, se evaluó la integridad del ácido nucleico total y del ARN en 1 µl utilizando el bioanalizador Agilent 2100. Las muestras se analizaron cualitativa y cuantitativamente. Se utilizó el análisis electroforético para comprobar la calidad del ARN extraído. Los resultados de la salida del Bioanalizador se analizaron mediante electroforesis en gel, como se muestra en la Figura 1 y la Figura 2. El archivo de electroforesis (Figura 1) se leyó comparando las bandas de cada muestra con las bandas representadas por los marcadores de tamaño en la escalera de ARN (mostrada en el primer carril del electroferograma) e identificando las bandas de ARN ribosómico 18S y 28S. El ARN ribosómico (ARNr) son las dos bandas grandes y prominentes alrededor del marcador de 2.000 nucleótidos en la escalera de normalización. Cualitativamente, un bandeo adecuado y unas intensidades de banda más oscuras indicaban que se disponía de abundante ácido nucleico intacto para análisis posteriores como la secuenciación de microarreglos, la reacción en cadena de la polimerasa (PCR), la secuenciación de ácidos nucleicos, el código de barras molecular o la captura de sondas. La Figura 2 muestra un ejemplo de electroferograma. El electroferograma es una representación gráfica de cada electroforesis con una cuantificación de la masa total de ARN, RIN y masa total de ARNr. El análisis por electroforesis en gel proporcionó información sobre el ácido nucleico total y la integridad del ARN. Se analizaron los diagramas de electroforesis y los electroferogramas de cada muestra para determinar si era apta para el análisis posterior. En un ejemplo, 120 muestras que se extrajeron utilizando los métodos descritos en la invención se analizaron para determinar la calidad del ARN, de forma que de las 120 muestras analizadas, 110 superaron la comprobación de calidad en términos de integridad del ARN y masa del ARN (Figura 3).

Los electroferogramas y las trazas de electroforesis se utilizaron para comparar las muestras que pasaron el control de calidad utilizando los métodos de extracción de ARN descritos en la invención en comparación con los métodos de extracción de ARN descritos en la bibliografía. En comparación con el protocolo descrito en la bibliografía, el método

de extracción de ácido nucleico descrito anteriormente aumentó el RIN medio a 4.88 (n=168). Además, el tiempo total necesario para completar el protocolo se redujo de 3 días a 5 horas. El método de extracción del ARN total aumentó significativamente el número de muestras aptas para el análisis del transcriptoma del 50 %, como se describe en la bibliografía, al 98.2 % en el protocolo descrito (n=168; p<0.0001), medido por la integridad y la masa del ARN ribosómico (Figura 4).

Ejemplo 4: Análisis de los transcritos de ARN

De las 454 muestras que fueron sometidas a extracción de ARN total, 399 de estas muestras fueron elegibles para el análisis del transcriptoma basado en la visualización directa de las bandas de ARN ribosómico utilizando un Bioanalizador Agilent. Posteriormente, se seleccionaron al azar 342 muestras de entre las 399 disponibles para someterlas a un análisis del transcriptoma completo utilizando un arreglo de Transcriptoma Humano 2.0 GeneChip™ de Affymetrix (Santa Clara, CA). De estas 342 muestras, 4 fallaron en la amplificación y se determinó que 8 pacientes con pólipos tenían pólipos benignos hiperplásicos mediante una biopsia posterior. Estas 12 muestras se eliminaron del análisis, lo que dio como resultado 330 muestras para el análisis final.

Se amplificaron aproximadamente 100 ng de ARN fecal libre de DNasa con el kit Ambio WT-pico con posterior hibridación al arreglo de Transcriptoma Humano 2.0 GeneChip™ de Affymetrix de acuerdo con el protocolo del fabricante. Todas las muestras se normalizaron utilizando la transformación del espacio de señal-análisis robusto de arreglos múltiples (SST-RMA) con Affymetrix Expression Console™.

De los 70.523 grupos de transcripción en el Microarreglo de Affymetrix, un subconjunto de 5,149 grupos de transcripción que corresponden a 3,977 genes fueron preseleccionados para evaluar la expresión diferencial. Esta selección inicial redujo la tasa de falsos descubrimientos y filtró genes sin función conocida en el desarrollo y la progresión del cáncer. Exploramos siete clasificadores de aprendizaje automático (Support Vector Machine, Gradient Boosting, Adaptive Boosting, Random Forest, Naive Bayes, Decision Tree y k-Nearest Neighbors) para la detección de neoplasias colorrectales.

Los 330 individuos se dividieron en un conjunto de entrenamiento de 265 individuos y un conjunto de prueba de 65 individuos. El conjunto de entrenamiento se utilizó para identificar los genes expresados diferencialmente y construir un modelo computacional, mientras que el conjunto de pruebas se utilizó para determinar la precisión de detección del modelo computacional. Se utilizó el paquete LIMMA estándar para identificar un subconjunto de grupos de transcripción de ARN que se expresaban de forma diferencial entre los individuos con adenomas precancerosos o CRC y los individuos sin hallazgos en una colonoscopia. Todos los biomarcadores se clasificaron de acuerdo con las puntuaciones de logaritmo de probabilidades y los 200 biomarcadores mejor clasificados (p<0.05) sirvieron como características para construir el modelo de aprendizaje automático (Figura 4). En la Figura 5A se resume un mapa de calor que describe la agrupación jerárquica de estos genes y en la Figura 5B se resumen las vías significativas asociadas a los genes expresados de forma diferencial. Había cinco grupos de genes asociados con los transcritos expresados diferencialmente en las tres poblaciones para segregar a los individuos con neoplasia colorrectal de los individuos sanos. Las vías expresadas diferencialmente mostraron un enriquecimiento para 22 vías canónicas comunes de GeneGO, en las que la gran mayoría (77.7 %) eran de regulación ascendente de la vía de los componentes celulares. Lo más notable fue el aumento de la respuesta celular al estrés, la respuesta al estrés y la unión del ARN. Se identificaron vías comunes asociadas a la enfermedad analizando los 200 grupos de transcripción más expresados diferencialmente con el software GAGE. Las vías asociadas a la enfermedad se muestran en la Figura 5B. Para el desarrollo del modelo se eligió el modelo de máquina de vectores de soporte (v-SVM) con núcleo RBF. La función de núcleo permite calcular la distancia entre individuos expandiendo las características a un espacio de mayor dimensión que no se calcula explícitamente. SVM encuentra el hiperplano de margen máximo que separa los grupos de etiquetas. El parámetro ν define el límite inferior de la fracción de individuos que se utilizan para determinar el margen máximo. El modelo SVM se entrenó utilizando los niveles de expresión de los 200 transcritos de todos los 265 individuos del conjunto de entrenamiento. Este algoritmo de biomarcadores de ARN multidiana se utilizó en los 65 individuos del conjunto de pruebas y detectó el 79 % (34 de 43) de todos los individuos que tuvieron resultados positivos en una colonoscopia de cribado. El algoritmo predijo correctamente el 95 % de los individuos con adenomas precancerosos de un tamaño comprendido entre 5 mm y 12 mm, el 57 % de los individuos con cáncer colorrectal en estadio I, el 75 % de los individuos con cáncer colorrectal en estadio II, el 66 % de los individuos con cáncer colorrectal en estadio II y el 83 % de los individuos con cáncer colorrectal en estadio IV. La sensibilidad del modelo para el cáncer colorrectal se correlacionó directamente con el tamaño, de forma que el 60 % de los tumores ≤ 3.5 cm se detectaron con precisión, mientras que el 83 % de los tumores ≥ 5.0 cm de tamaño se detectaron con precisión. El modelo alcanzó una especificidad del 59 %, por lo que 13 de los 22 individuos con resultados negativos en una colonoscopia fueron identificados correctamente por el modelo. En conjunto, el modelo alcanzó una sensibilidad para las neoplasias colorrectales del 79 %, una sensibilidad para los adenomas precancerosos del 95 % y una especificidad del 59 %.

Las réplicas tecnológicas y biológicas mostraron una alta concordancia para los niveles de luminiscencia de todos los genes (Figura 10). Las réplicas técnicas perfectas, en las que el ARN se aisló y evaluó el mismo día, mostraron el mayor nivel de concordancias ($R^2 > 0.990$) (Figura 10A). Las réplicas tecnológicas separadas por el tiempo también mostraron un alto nivel de concordancia ($R^2 > 0.989$). Estas réplicas se obtuvieron extrayendo el ARN de ambas réplicas el mismo día; sin embargo, una réplica se analizó el día 0 y una segunda réplica se analizó 6 meses después (Figura 10B). Por último, las réplicas biológicas también mostraron un alto nivel de concordancias ($R^2 > 0.986$). Estas réplicas

biológicas se desarrollaron tomando muestras de diferentes segmentos de una deposición procedente del mismo individuo (Figura 10C). El ARN se extrajo por separado y se analizó en días diferentes.

Al analizar sólo 5,149 genes que han sido implicados en el desarrollo o progresión del cáncer colorrectal, un subconjunto del transcriptoma completo, los valores R-cuadrado ajustados mejoraron para las tres cohortes de réplicas. Estas réplicas mostraron niveles de concordancia comparables a los descritos anteriormente (Figura 11A-C).

Ejemplo 5: Obtención de muestras de heces de animales

Colección de heces: Las muestras fueron recogidas localmente en St. Louis, MO, por propietarios de perros y gatos. Una vez que los gatos defecaban en una caja de arena, se pedía a los propietarios de los felinos que transfirieran la muestra a un tubo cónico de 50 ml y que guardaran el tubo a -20 °C. Cuando los perros defecaban al aire libre, se pedía a los propietarios de los caninos que transfirieran la muestra a un tubo cónico de 50 ml y lo guardaran a -20 °C. Una semana después de la recogida, las muestras se recogían y se transferían manualmente a BioGenerator Labs (Saint Louis, MO), donde se almacenaban en un congelador a -80 °C hasta su extracción.

Tipos de muestras: Todas las muestras de heces se obtuvieron de animales sanos de distintas edades, razas y sexos. se recogieron 4 muestras de 4 gatos diferentes y otras 4 muestras de 4 perros diferentes, lo que dio como resultado 8 muestras agregadas. Un gato había sido diagnosticado recientemente de tiña, pero todos los demás animales eran asintomáticos y no habían mostrado signos de malestar gastrointestinal en los últimos 30 días.

Ejemplo 6: Extracción de ácido nucleico animal

Extracción de ácido nucleico total: Cada muestra de heces se colocó en un tubo cónico de 50 ml. Se añadieron aproximadamente 1.000-25.000 mg de heces a cada tubo. Se añadieron otros 20-40 ml de solución a cada tubo. Esta solución contenía una mezcla de solución salina equilibrada de Hanks (HBSS) (Sigma-Aldrich) con 0.05 % de Tween-20 (Sigma-Aldrich) y 0.0002 % de inhibidor de RNasa (Sigma-Aldrich). Las heces se suspendieron en la solución y se rotaron a aproximadamente 0-10 °C durante 0-10 minutos. La solución se centrifugó a 1000 rpm a 4 °C durante 10 minutos y se desechó el sobrenadante. Se añadieron aproximadamente 4-10 ml de tampón de lisis EasyMag® (bioMerieux) al gránulo y el gránulo se volvió a suspender en la solución. La solución se centrifugó a 2500-3500 rpm a 20-25 °C durante 10-15 minutos. Durante la centrifugación diferencial, la solución se separó en tres capas. La capa inferior incluía restos celulares sólidos, la capa intermedia era una capa hidrófila enriquecida en ácido nucleico humano y la capa superior era una capa lipídica hidrófoba. Las dos capas superiores se transfirieron a un nuevo tubo cónico de 15 ml y la solución se centrifugó de nuevo a 2500 rpm a 20-25 °C durante 10 minutos. El resultado de esta etapa de centrifugación fue la separación en tres capas: la capa inferior eran restos celulares sólidos, la capa intermedia era una capa hidrófila enriquecida en ácido nucleico humano y la capa superior era una capa lipídica hidrófoba. Para filtrar los residuos grandes de la solución, se colocó una punta de pipeta de 20 ul en una punta de pipeta de 1 ml y se pipetearon 2 ml de la capa hidrófila del tubo de 15 ml y se transfirieron a un cartucho desechable EasyMag® (bioMerieux). Además, se añadieron 60 ul de sílice magnética EasyMag® (bioMerieux) al cartucho. Las perlas se mezclaron en la solución durante 0.5-1 minutos utilizando una pipeta. Los ácidos nucleicos, que estaban unidos a las perlas, se eluyeron en una solución tampón utilizando el Protocolo A Específico de acuerdo con las instrucciones del fabricante. El volumen de los ácidos nucleicos eluidos fue de 70 ul. Esta solución de ácido nucleico se pipeteó en un tubo de 1.5 ml y se colocó en hielo. Los mismos cartuchos desechables EasyMag® (bioMerieux) que se utilizaron en la etapa anterior se volvieron a cargar con 2 ml adicionales de la capa hidrófila de la misma solución en el tubo de 15 ml utilizado anteriormente utilizando la misma técnica para filtrar los residuos grandes. Se añadieron 20 ul adicionales de sílice magnética EasyMag® (bioMerieux) al cartucho. Las perlas se mezclaron en la solución durante 0.5-1 minutos utilizando una pipeta. Como se ha descrito anteriormente, los ácidos nucleicos, que estaban unidos a las perlas, se eluyeron en una solución tampón utilizando el Protocolo A Específico de acuerdo con las instrucciones del fabricante. El volumen de los ácidos nucleicos eluidos fue de 70 ul. Esta solución de ácido nucleico se pipeteó en el tubo original de 1.5 ml que ya contenía el primer eluido de 70 ul y la solución combinada se colocó en hielo.

Tratamiento con DNasa: La solución de 140 ul se trató con Baseline-Zero-DNasa (Epicenter) a 35-40 °C durante 20-40 minutos. Se añadió una alícuota de 1-2 ml de tampón de lisis EasyMag® a la solución tratada con DNasa y se transfirió la muestra a un nuevo cartucho EasyMag® desechable. Se añadió toda la solución al nuevo cartucho junto con 60 ul de sílice magnética EasyMag®. Los ácidos nucleicos, que estaban unidos a las perlas, se eluyeron en una solución tampón utilizando el protocolo EasyMag® Generic de acuerdo con las instrucciones del fabricante. El volumen de los ácidos nucleicos eluidos fue de 25 ul. Esta solución de ácido nucleico se pipeteó en un tubo de 1.5 ml y se almacenó a 0-6 °C.

Ejemplo 7: Medición de los niveles de ácidos nucleicos en muestras de heces de animales

Resultados de la extracción: Utilizando las muestras extraídas anteriormente, se evaluó la integridad del ácido nucleico total y del ARN de 1 ul de cada una utilizando el bioanalizador Agilent 2100. Las muestras se analizaron cualitativa y cuantitativamente. Los resultados de la salida del Bioanalizador se analizaron mediante electroforesis en gel, como se muestra en la Figura 7. El panel de la Figura 7 muestra las trazas del bioanalizador de ARN de 1 ul de 8 muestras que se analizaron con los kits Agilent RNA 6000 Nano. El análisis de electroforesis en gel proporcionó información sobre el ácido nucleico total y la integridad y masa del ARN. El archivo de electroforesis se leyó comparando las bandas de

cada muestra con las bandas representadas por los marcadores de tamaño en la escalera de ARN (mostrada en el primer carril del electroferograma) e identificando las bandas de ARN ribosómico eucariota 18S y 28S. Cualitativamente, un bandeo adecuado y unas intensidades de banda más oscuras indicaban que se disponía de abundante ácido nucleico intacto para análisis posteriores como la secuenciación de microarreglos, la reacción en cadena de la polimerasa (PCR), la secuenciación de ácidos nucleicos, el código de barras molecular o la captura de sondas. La calidad de la muestra también se analizó cualitativamente mediante el número de integridad del ARN. La Figura 7B muestra el valor de integridad del ARN para cada una de las mismas 8 muestras que se analizaron con los kits Agilent RNA 6000 Nano. La masa de la muestra también se analizó cuantitativamente. La Figura 7C muestra la concentración estimada de ARN eucariota (ug/ul) de las mismas 8 muestras que se analizaron con los kits Agilent RNA 6000 Nano. La concentración de ARN eucariota se estimó utilizando el área bajo la curva de las trazas del BioAnalyzer. Como se muestra en la Figura 7B-C, el RIN promedio para todas las muestras fue de 4.2 y la masa eucariota promedio para todas las muestras fue de 61.4 ng/ul. Las mediciones cuantitativas y cualitativas indicaron que todas las muestras eran aptas para el análisis de expresión de biomarcadores (n=8).

Ejemplo 8: Análisis de transcritos de ARN en muestras de heces de animales

Todas las muestras caninas elegibles para el análisis de expresión de biomarcadores (n=4) se evaluaron mediante RT-qPCR para estas cuatro muestras caninas, se diseñaron 5 cebadores para evaluar porciones del genoma canino. Se diseñaron dos cebadores para actuar como control positivo evaluando la cadena pesada de IgM en células T caninas. Se diseñaron tres cebadores adicionales para evaluar dos reordenamientos distintos que pueden producirse en precursores linfoides caninos para detectar ARN específico de células T. Cada reacción de RT-PCR contenía 10 ul de 2X Master Mix (New England Biolabs), 1 ul de 20X Enzyme Mix (New England BioLabs), 0.8 ul de 10 uM de cebador directo (Eurofins Genomics), 0.8 ul de 10 uM de cebador inverso (Eurofins Genomics), 1ul de 20X SYBR Green I, 2 ul de ARN y 4.4 ul de H₂O de calidad para biología molecular (Thermo Scientific). Los ácidos nucleicos se amplificaron utilizando una máquina qPCR QuantStudio5 de Applied Biosystems. El protocolo del termociclador fue el siguiente: 25 minutos a 55 °C para la transcripción inversa del ARN en ADN, inactivación de la transcriptasa inversa a 95 °C durante 1 minuto, 60 ciclos de 10 segundos a 95 °C y 45 segundos a 60 °C para la amplificación y recogida de la señal, y a continuación una curva de fusión de 70 °C-95 °C. De las 4 muestras caninas sometidas a extracción de ARN total, el 75 % de las muestras mostraron amplificación del control positivo (C-mu IgM) (Figura 8A). De los individuos que mostraron amplificación del control positivo, el 100 % (n=3) mostraron amplificación de ambos transcritos de ARN específicos de células T. Uno de los transcritos específicos de células T amplificó a 35 ciclos y el otro transcrito específico de células T amplificó a 50 ciclos (Figura 8B).

Todas las muestras felinas elegibles para el análisis de expresión de biomarcadores (n=5) se evaluaron mediante RT-qPCR. Para estas cuatro muestras felinas, se diseñaron dos cebadores que actuaran como control positivo. Estas sondas evaluaron la Actina-B felina para determinar si había ARN felino presente en el extracto de ácido nucleico. Cada reacción contenía: 10 ul de 2X Master Mix (New England BioLabs), 1 ul de 20X Enzyme Mix (New England BioLabs), 0.8 ul de 10 uM de cebador directo (Eurofins Genomics), 0.8 ul de 10 uM de cebador inverso (Eurofins Genomics), 1 ul de 20X SYBR Green I, 2 ul de ARN y 4.4 ul de H₂O de calidad para biología molecular (Thermo Scientific). La solución se amplificó utilizando una máquina qPCR QuantStudio5 de Applied Biosystems. El protocolo del termociclador fue el siguiente: 25 minutos a 55 °C para la transcripción inversa del ARN en ADN, inactivación de la transcriptasa inversa a 95 °C durante 1 minuto, 60 ciclos de 10 segundos a 95 °C y luego 45 segundos a 60 °C para la amplificación y la recogida de la señal. De las 4 muestras felinas sometidas a extracción de ARN total, el 75 % de las muestras (n=4) mostraron amplificación del control positivo (Actina-B felina) (Figura 9).

Ejemplo 9: Análisis de los transcritos de las células B caninas

Todas las muestras caninas elegibles para el análisis de expresión de biomarcadores (n=4) se evaluaron mediante RT-qPCR. Para estas 4 muestras caninas, se diseñaron 6 conjuntos de cebadores para evaluar porciones del genoma canino relacionadas con los linfocitos (Figura 15A). De los 6 conjuntos de cebadores diseñados, uno se utilizó para evaluar dos reordenamientos distintos que pueden producirse en precursores linfoides caninos para detectar ARN específico de células B. Cada reacción contenía 10 ul de 2X Master Mix (New England Biolabs), 1 ul de 20X Enzyme Mix (New England BioLabs), 0.8 ul de 10 uM de cebador directo (Eurofins Genomics), 0.8 ul de 10 uM de cebador inverso (Eurofins Genomics), 1 ul de 20X SYBR Green I, 2 ul de ARN y 4.4 ul de H₂O de calidad para biología molecular (Thermo Scientific). La solución se amplificó utilizando una máquina qPCR QuantStudio5 de Applied Biosystems. El protocolo del termociclador fue el siguiente: 25 minutos a 55 °C para transcribir inversamente el ARN en ADN, inactivación de la transcriptasa inversa a 95 °C durante 1 minuto, 60 ciclos de 10 segundos a 95 °C y 45 segundos a 60 °C para la amplificación y la recogida de la señal, y luego una curva de fusión de 70 °C-95 °C. En la Figura 16A se muestran ejemplos de estas reacciones de amplificación. Hubo 2 reacciones RT-qPCR específicas de células B para cada muestra (n=8) y de estas reacciones, el 87.5 % (7 de 8) mostraron amplificación de transcritos relacionados con células B. Los transcritos específicos de células B se amplificaron entre los ciclos 30 y 35 (Figura 16A).

Ejemplo 10: Extracción de réplicas de muestras felinas

Recogida de heces: Las muestras fueron recogidas localmente en St. Louis, MO por propietarios de gatos. Una vez que los gatos defecaban en una caja de arena, se pedía a los propietarios de los felinos que transfirieran la muestra a un tubo cónico de 50 ml y almacenaran el tubo a -20 °C. En el plazo de una semana desde la producción, se recogieron

las muestras y se transfirieron manualmente a BioGenerator Labs (Saint Louis, MO), donde se almacenaron en un congelador a -80 °C hasta su extracción.

Tipos de muestras: Todas las muestras de heces se obtuvieron de animales sanos de distintas edades, razas y géneros. Se recogieron 8 muestras de 4 gatos diferentes. Para algunos gatos, se recogieron hasta tres réplicas biológicas. Nos referimos a una réplica biológica como una muestra de heces del mismo gato recogida de deposiciones separadas. Un gato había sido diagnosticado recientemente de tiña, pero todos los demás animales eran asintomáticos y no habían mostrado signos de malestar gastrointestinal en los últimos 30 días.

Extracción de ácido nucleico total: Cada muestra de heces se colocó en un tubo cónico de 50 ml. Se añadieron aproximadamente 1.000-25.000 mg de heces a cada tubo. Se añadieron otros 20-40 ml de solución a cada tubo. Esta solución contenía una mezcla de solución salina equilibrada de Hanks (HBSS) (Sigma-Aldrich) con 0.05 % de Tween-20 (Sigma-Aldrich) y 0.0002 % de inhibidor de RNasa (Sigma-Aldrich). Las heces se suspendieron en la solución y se rotaron a aproximadamente 0-10 °C durante 0-10 minutos. La solución se centrifugó a 1000 rpm a 4 °C durante 10 minutos y se desechó el sobrenadante. Se añadieron aproximadamente 4-10 ml de tampón de lisis EasyMag® (bioMerieux) al gránulo y éste se volvió a suspender en la solución. La solución se centrifugó a 2500-3500 rpm a 20-25 °C durante 10-15 minutos. Durante la centrifugación diferencial, la solución se separó en tres capas. La capa inferior incluía restos celulares sólidos, la capa intermedia era una capa hidrófila enriquecida en ácido nucleico humano y la capa superior era una capa lipídica hidrófoba. Las dos capas superiores se transfirieron a un nuevo tubo cónico de 15 ml y la solución se centrifugó de nuevo a 2500 rpm a 20-25 °C durante 10 minutos. El resultado de esta etapa de centrifugación fue la separación en tres capas: la capa inferior eran restos celulares sólidos, la capa intermedia era una capa hidrófila enriquecida en ácido nucleico humano y la capa superior era una capa lipídica hidrófoba. Para filtrar los residuos grandes de la solución, se colocó una punta de pipeta de 20 µl en una punta de pipeta de 1 ml y se pipetearon 2 ml de la capa hidrófila del tubo de 15 ml y se transfirieron a un cartucho desechable EasyMag® (bioMerieux). Además, se añadieron 60 µl de sílice magnética EasyMag® (bioMerieux) al cartucho. Las perlas se mezclaron en la solución durante 0.5-1 minutos utilizando una pipeta. Los ácidos nucleicos, que estaban unidos a las perlas, se eluyeron en una solución tampón utilizando el Protocolo A Específico de acuerdo con las instrucciones del fabricante. El volumen de los ácidos nucleicos eluidos fue de 70 µl. Esta solución de ácido nucleico se pipeteó en un tubo de 1.5 ml y se colocó en hielo. Los mismos cartuchos desechables EasyMag® (bioMerieux) que se utilizaron en la etapa anterior se volvieron a cargar con 2 ml adicionales de la capa hidrófila de la misma solución en el tubo de 15 ml utilizado anteriormente utilizando la misma técnica para filtrar los residuos grandes. Se añadieron 20 µl adicionales de sílice magnética EasyMag® (bioMerieux) al cartucho. Las perlas se mezclaron en la solución durante 0.5-1 minutos utilizando una pipeta. Como se ha descrito anteriormente, los ácidos nucleicos, que estaban unidos a las perlas, se eluyeron en una solución tampón utilizando el Protocolo A Específico de acuerdo con las instrucciones del fabricante. El volumen de los ácidos nucleicos eluidos fue de 70 µl. Esta solución de ácido nucleico se pipeteó en el tubo original de 1.5 ml que ya contenía el primer eluido de 70 µl y la solución combinada se colocó en hielo.

Tratamiento con DNasa: La solución de 140 µl se trató con Baseline-Zero-DNasa (Epicenter) a 35-40 °C durante 20-40 minutos. Se añadió una alícuota de 1-2 ml de tampón de lisis EasyMag® a la solución tratada con DNasa y se transfirió la muestra a un nuevo cartucho EasyMag® desechable. Se añadió toda la solución al nuevo cartucho junto con 60 µl de sílice magnética EasyMag®. Los ácidos nucleicos, que estaban unidos a las perlas, se eluyeron en una solución tampón utilizando el protocolo EasyMag® Generic de acuerdo con las instrucciones del fabricante. El volumen de los ácidos nucleicos eluidos fue de 25 µl. Esta solución de ácido nucleico se pipeteó en un tubo de 5 ml y se almacenó a 0-6 °C.

Ejemplo 11: Medición de los niveles de ácido nucleico en réplicas de muestras felinas

Resultados de la extracción: Utilizando las muestras extraídas anteriormente, se evaluó la integridad del ácido nucleico total y del ARN de 1 µl de cada una utilizando el bioanalizador Agilent 2100. Las muestras se analizaron cualitativa y cuantitativamente. Los resultados de la salida del Bioanalizador se analizaron mediante electroforesis en gel, como se muestra en la Figura 14. La Figura 14 muestra las trazas del bioanalizador de ARN de 1 µl de 8 muestras que se analizaron con los kits Agilent RNA 6000 Nano. Esto incluía 4 gatos individuales. El análisis de electroforesis en gel proporcionó información sobre el ácido nucleico total y la integridad y masa del ARN. El archivo de electroforesis se leyó comparando las bandas de cada muestra con las bandas representadas por los marcadores de tamaño en la escalera de ARN (mostrada en el primer carril del electroferograma) e identificando las bandas de ARN ribosómico eucariota 18S y 28S. Cualitativamente, un bandeo adecuado y unas intensidades de banda más oscuras indicaban que se disponía de abundante ácido nucleico intacto para análisis posteriores como la secuenciación de microarreglos, la reacción en cadena de la polimerasa (PCR), la secuenciación de ácidos nucleicos, el código de barras molecular o la captura de sondas. el 100 % de las muestras eran aptas para el análisis de acuerdo con el número de integridad del ARN y la cantidad de masa eucariota. La concentración eucariota se estimó a partir del área bajo las bandas de ARN ribosómico eucariota 18S y 28S de las trazas del BioAnalyzer. La Figura 14 muestra réplicas biológicas para 4 gatos individuales. Se observaron bandas similares entre réplicas biológicas y bandas distintas para gatos individuales (Figura 14).

Ejemplo 12: Análisis de transcritos de ARN en réplicas de muestras felinas

Todas las muestras felinas individuales elegibles para el análisis de expresión de biomarcadores (n=5) se evaluaron mediante RT-qPCR. Se diseñó un total de 4 conjuntos de cebadores para evaluar los transcritos relacionados con los linfocitos y los controles relacionados con las células epiteliales (Figura 15B). Se diseñó un conjunto de cebadores para evaluar los reordenamientos que pueden producirse en los precursores linfoides felinos para detectar ARN específico de células T. Cada reacción contenía 10 µl de 2X Master Mix (New England BioLabs), 1 µl de 20X Enzyme Mix (New England BioLabs), 0.8 µl de 10 µM de cebador directo (Eurofins Genomics), 0.8 µl de 10 µM de cebador inverso (Eurofins Genomics), 1 µl de 20X SYBR Green I, 2 µl de ARN y 4.4 µl de H₂O de calidad para biología molecular (Thermo Scientific). La solución se amplificó utilizando una máquina qPCR QuantStudio5 de Applied Biosystems. El protocolo del termociclador fue el siguiente: 25 minutos a 55 °C para la transcripción inversa del ARN en ADN, inactivación de la transcriptasa inversa a 95 °C durante 1 minuto, 60 ciclos de 10 segundos a 95°C y luego 45 segundos a 60 °C para la amplificación y la recogida de la señal. En la Figura 16B se muestran ejemplos de estos eventos de amplificación. De las cinco muestras felinas que se analizaron en busca de reordenamientos selectos de células T, en un total de 12 reacciones, el 75 % de las reacciones mostraron amplificación de los transcritos relacionados con las células T. Todos los reordenamientos de células T analizados mostraron amplificación en al menos un gato (Figura 16B).

Ejemplo 13: Análisis NanoString

Para analizar más a fondo los genes asociados con el cáncer colorrectal humano y los adenomas precancerosos, se sometieron 70 muestras humanas adicionales al método de extracción de ARN descrito anteriormente. El ARN extraído se analizó mediante el sistema de análisis NanoString nCounter, que utiliza una tecnología digital de código de barras codificado por colores que se basa en la medición multiplexada directa de la expresión génica. De las 70 muestras, 48 se analizaron utilizando el panel Counter® PanCancer Pathways, que incluye 770 genes de 13 vías canónicas asociadas al cáncer, que incluyen: MAPK, STAT, PI3K, RAS, ciclo celular, apoptosis, Hedgehog, Wnt, control del daño en el ADN, regulación transcripcional, modificación de la cromatina y TGF-β. Las 22 muestras restantes se analizaron utilizando el Panel de Progresión PanCáncer nCounter®, que incluye 770 genes de cada etapa del proceso de progresión del cáncer, que incluye: angiogénesis, remodelación de la matriz extracelular (ECM), transición epitelio-mesénquima (EMT) y metástasis. Estas 22 muestras también fueron evaluadas para 10 genes “añadidos como control”), que incluyen ACTB, B2M, BMP3, CD274, CD8A, GAPDH, HPRT1, N-BLR, NDRG4 y RNU2-1.

Además de la lista de genes identificados en la Figura 6 (Panel A), los 200 principales genes expresados diferencialmente, también analizamos genes adicionales que estaban asociados con el cáncer colorrectal y los adenomas precancerosos. También analizamos todos los genes que se identificaron como altamente expresados o diferencialmente expresados en el sistema de análisis NanoString nCounter. Este análisis incluyó el uso del panel nCounter® PanCancer Pathways, el panel nCounter® PanCancer Progression y los genes “añadidos como control”. Nosotros utilizamos esta información para desarrollar una firma de biomarcadores de ARN de 400 genes para el cáncer colorrectal y los adenomas precancerosos (Figura 13 (Panel B)).

REIVINDICACIONES

1. Un método de aislamiento de ácido nucleico eucariota a partir de una muestra de heces, el método comprende:
 - a) mezclar la muestra con un tampón, un tensioactivo y un inhibidor de ribonucleasas para formar una suspensión;
 - b) separar la suspensión en una porción enriquecida para células eucariotas y una porción enriquecida para células bacterianas y retener la porción enriquecida para células eucariotas;
 - c) añadir un agente caotrópico y, opcionalmente, un tensioactivo a la porción enriquecida en células eucariotas para formar un lisado;
 - d) fraccionar el lisado en una capa de restos celulares, una capa que comprenda ácidos nucleicos eucariotas y una capa lipídica; y
 - e) recoger la capa que comprenda los ácidos nucleicos eucariotas y, opcionalmente, la capa lipídica.
2. El método de la reivindicación 1, en el que la muestra de heces es:
una muestra de heces humanas
3. El método de la reivindicación 1, en el que la muestra de heces es:
una muestra de heces de un animal no humano, opcionalmente en la que el animal no humano se selecciona del grupo que consiste en perro, gato, primate no humano, rumiante, úrsido, équido, cerdo, oveja, cabra, camélido, búfalo, ciervo, ciervo americano, alce, mustélido, conejo, cobaya, hámster, rata, ratón, paquidermo, rinoceronte y chinchilla.
4. El método de cualquier reivindicación precedente, en el que la etapa de separación comprende:
centrifugación o un método basado en columnas o utiliza la filtración diferencial.
5. El método de cualquier reivindicación precedente, en el que la etapa de separación comprende:
utilizar sondas diana que se unen a células eucariotas.
6. El método de cualquier reivindicación precedente, en el que la etapa de fraccionamiento comprende:
centrifugación o utiliza la filtración diferencial.
7. El método de cualquier reivindicación precedente, en el que la etapa de fraccionamiento comprende:
utilizar sondas diana que se unen específicamente al ácido nucleico eucariota.
8. El método de cualquier reivindicación precedente, en el que:
las etapas a y b se repiten una, dos, tres, cuatro o más veces.
9. El método de cualquier reivindicación precedente, en el que:
las etapas d y e se repiten dos o más veces.
10. El método de cualquier reivindicación precedente, en el que:
la capa que contiene ácidos nucleicos eucariotas se recoge usando una punta de pequeño calibre.
11. El método de cualquier reivindicación precedente, en el que:
el ácido nucleico comprende ADN, ARN, ARN total, ARNm, ARNt, ARNr, ARNnc, ARNsm o ARNsno, o una combinación de cualquiera de ADN, ARN, ARN total, ARNm, ARNt, ARNr, ARNnc, ARNsm o ARNsno.
12. El método de cualquier reivindicación precedente, en el que:
el método comprende además:
 - i) extraer los ácidos nucleicos eucariotas de la capa recogida que comprende ácidos nucleicos eucariotas; y/o
 - ii) proporcionar una muestra de heces.
13. El método de la reivindicación 12(i), en el que: el método de extracción es un método basado en partículas magnéticas, en columnas, en filtros, en perlas o en disolventes orgánicos.
14. Un método de detección de un biomarcador eucariota en una muestra de heces, el método comprende:

realizar el método de cualquiera de las reivindicaciones 1-13, que comprende además analizar el ácido nucleico extraído mediante secuenciación de microarreglos, código de barras molecular, captura de sondas, reacción en cadena de la polimerasa (PCR), ddPCR, RT-PCR, RT-qPCR, o secuenciación de ácidos nucleicos; opcionalmente en donde:

- 5 el biomarcador eucariota es i) seleccionado entre los biomarcadores listados en la Figura 6 o la Figura 13 o ii) es un marcador de células B, un marcador de células T o una inmunoglobulina.

Figura 1: Ejecución de un archivo de electroforesis

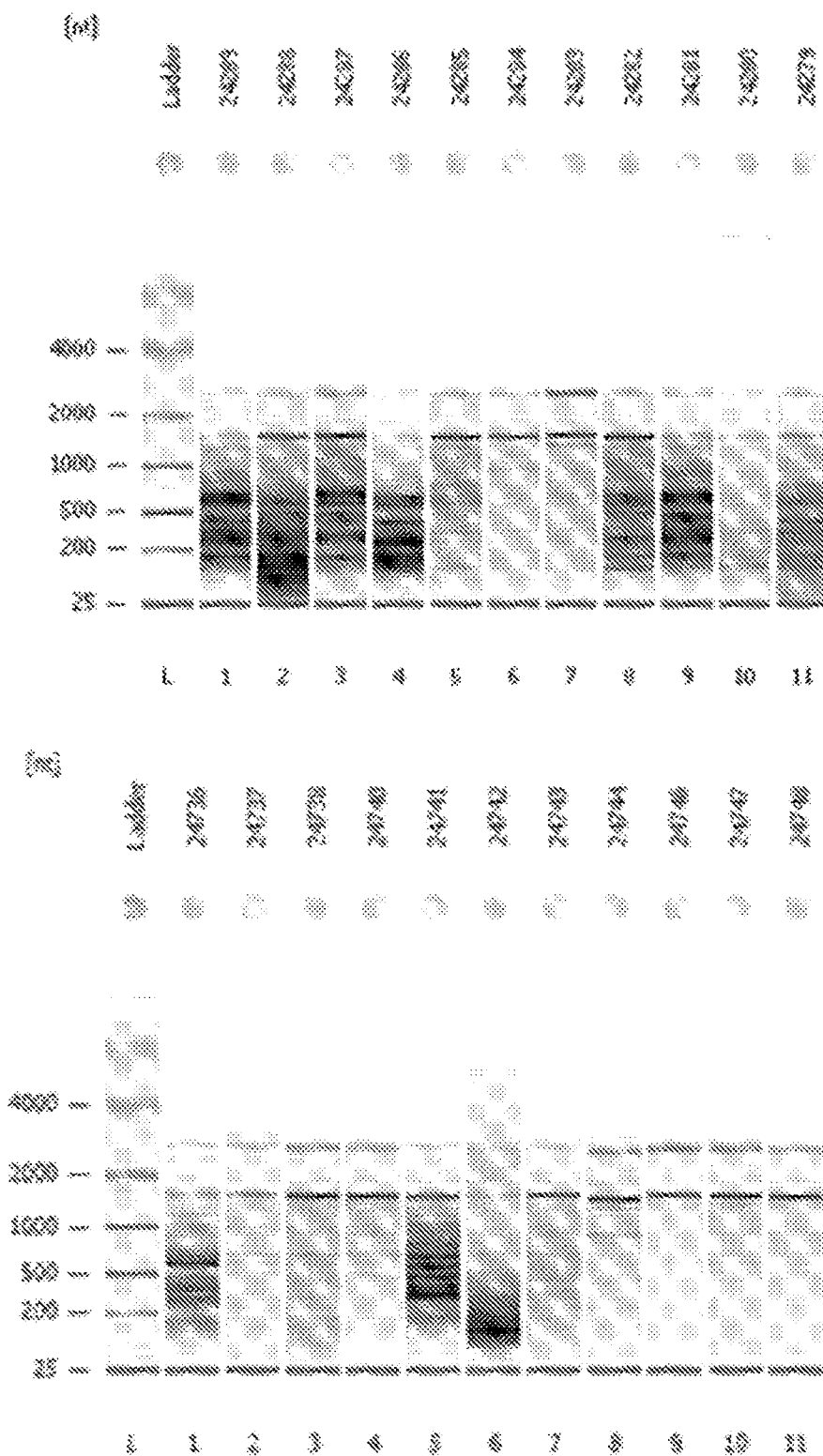
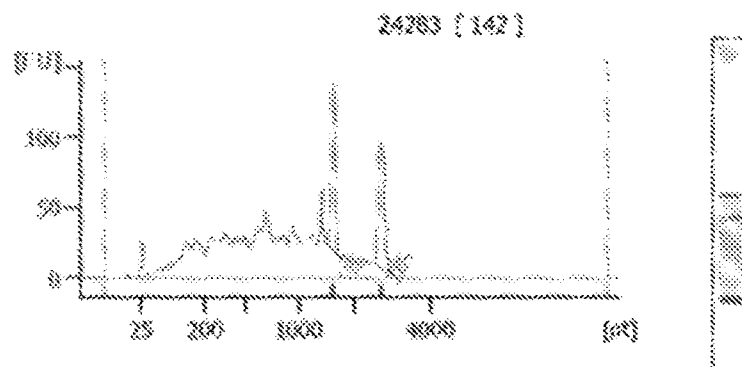


Figura 2: Electroferograma de muestra**Resultados generales para muestra 7:****24283**

Área de ARN: 1,703.0
 Concentración de ARN: 4,087 pg/μl
 Relación de ARN: 1.0
 Número de integridad del ARN (RIN): 5.1 (8.82.08)
 Color de marcado de resultado:
 Etiqueta de marcado de resultado: 5.1

Tabla de fragmento para la muestra 7:**24283**

| Nombre | Tamaño de inicio [nt] | Tamaño final [nt] | Área | % de área total |
|--------|-----------------------|-------------------|-------|-----------------|
| 159 | 1,424 | 1,706 | 921.3 | 7.1 |
| 285 | 2,443 | 2,900 | 120.8 | 7.1 |

Figura 3: Resumen de extracción

| | |
|---|--------------|
| Número de muestras en estudio | 120 |
| Numero de muestras que pasan QC | 110 |
| Numero de muestras que fallan QC | 10 |
| % de muestras que pasan QC | 91.7% |

Figura 4: Método de extracción para aislar biomarcadores de ARNm humano en muestras de Heces

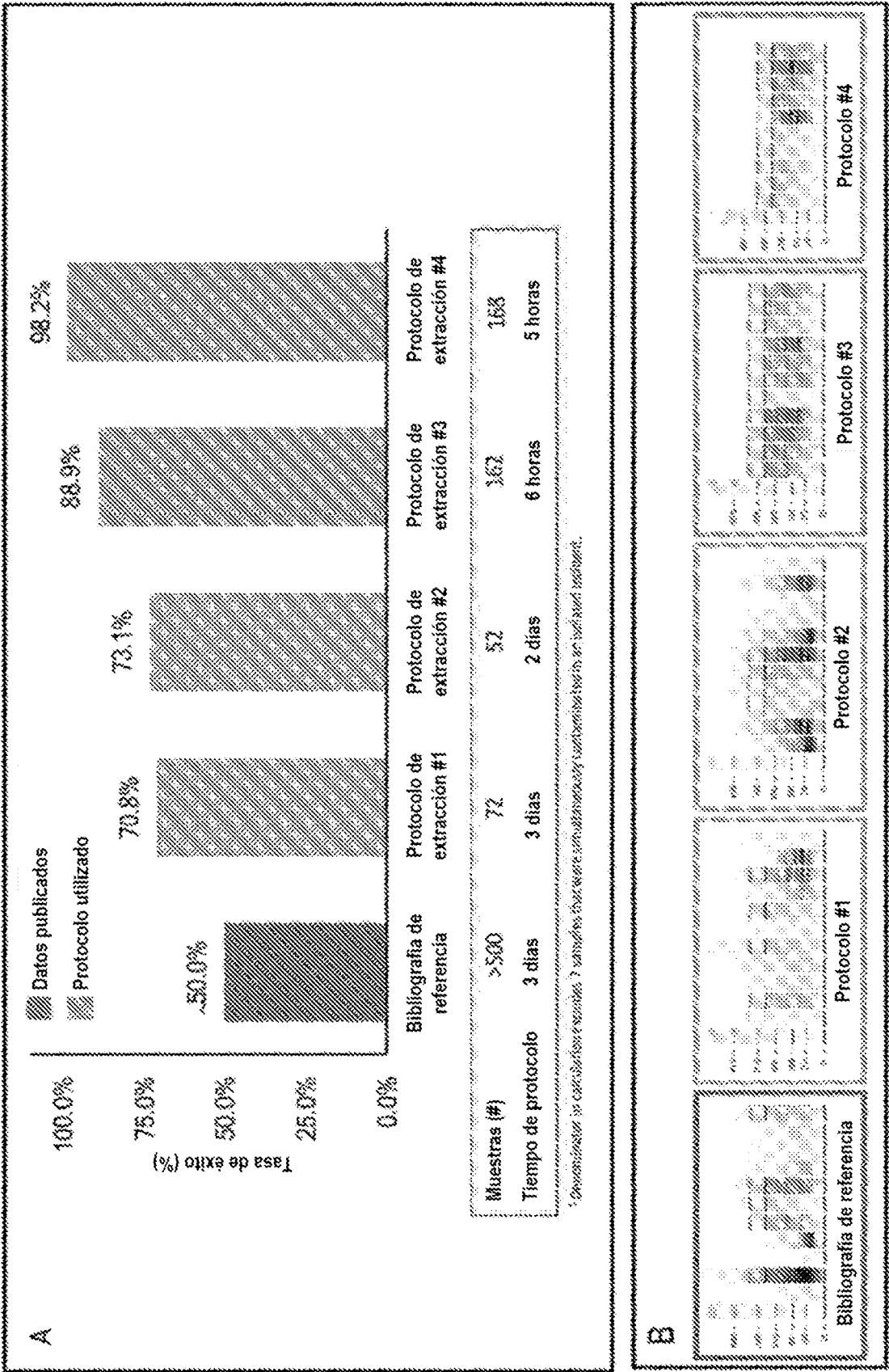


Figura 5: Anotaciones génicas diferencialmente expresadas

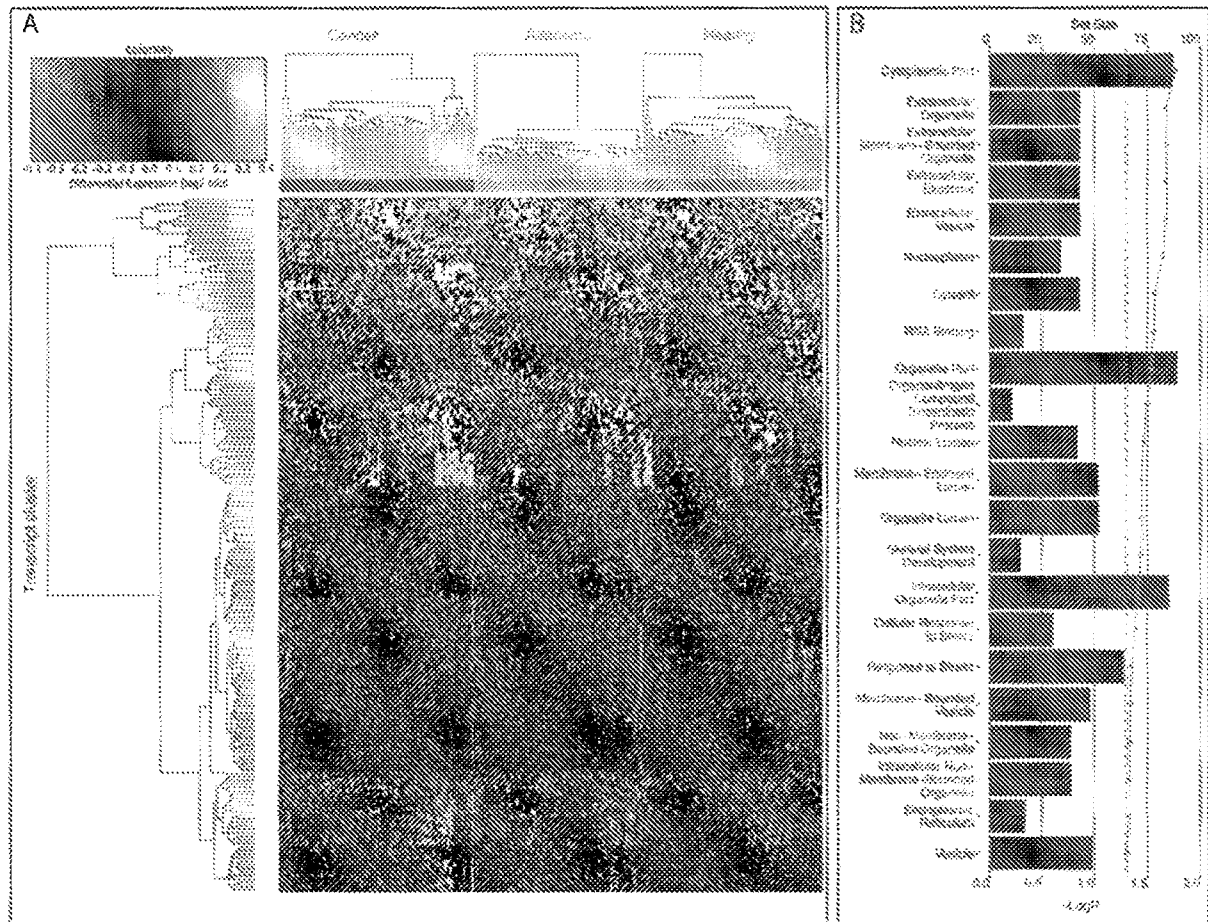


Figura 6: Genes diferencialmente expresados. Lista de nombres de genes basados en nomenclatura **HUGO** que están diferencialmente expresados a ($p < 0.05$) en neoplasmas colorrectales.

| | | | | |
|----------|----------|----------|-----------|----------|
| COMP | GOLPH3 | FGF6 | CYP1B1 | MAGED4 |
| SACS | KLK4 | TTLL3 | MYB | CHD7 |
| PPP3CB | JAG2 | NCRUPAR | LTA | ID2 |
| ERCC2 | GLRX3 | TFDP1 | CLEC3B | SPARCL1 |
| TAC1 | TNFSF9 | GISH | KISS1R | TUBB3 |
| LRP5 | AIM2 | ASCL1 | FCGR3A | CD151 |
| EPOR | KALRN | PTGER2 | PML | IFNAR1 |
| PARP11 | CCNB1 | POLD1 | TMPRSS4 | FRMPD2 |
| CLDN4 | BVES | SLC25A46 | PLCD1 | OLR1 |
| SCGB2A1 | ERBB3 | CDH22 | MSH5 | GSTA4 |
| IL23A | PTPRS | TLR2 | CDC20 | B4GALT3 |
| GNB2 | MDK | CSRNP3 | RPIA | HMGCR |
| PRKCB | VTCN1 | BR3 | NCK1 | CTX1 |
| HHEX | ADCY1 | ARID1B | ARL4C | HOXA5 |
| FDPS | PER3 | CITED1 | DHX16 | RPRM |
| GNG8 | THBD | MNX1 | STK11 | RALGDS |
| TMEM158 | IL20RB | PYCR1 | HIST1H2BM | WDR49 |
| ROR1 | TENM1 | CEBPA | PLA2G2A | S100A14 |
| CD276 | TSPEAR | SPINK6 | CDC25C | COL3A1 |
| MIR129-2 | SFRP5 | PEA15 | HK1 | MAPK3 |
| TLR3 | HLA-DQB1 | ZNF135 | NFIB | WNT3A |
| PTGES2 | CSF3R | MYNN | CDHR5 | TCN2 |
| WNT10A | SALL3 | BMP3 | BDNF | BOLL |
| EPHA7 | MUC6 | HLA-DRB1 | CUX1 | MCM5 |
| MAP2K3 | NR4A3 | SHROOM2 | ACY1 | PRMT8 |
| SHC2 | FASN | HK2 | TNS4 | RFX1 |
| NGF | DNAJC14 | CASC2 | ATP1B1 | MDC1 |
| ATP11A | TMEFF2 | CXCL13 | HSPB6 | GDNF |
| SLC12A5 | CSF1R | CCDC129 | TWIST1 | ROBO1 |
| ENO1 | COMT | DPP6 | IGFBP3 | PKMYT1 |
| CD8A | UBD | FNTB | MIR148B | CABLES1 |
| CCNG2 | SALL4 | WNT9A | PTPRU | KDM5C |
| MGST1 | LGR5 | AMD1 | EHMT2 | NID1 |
| ADCY5 | UHRF1 | FOFT1 | EMP1 | SMC1A |
| DHRS9 | MIR223 | PGGT1B | FNTA | CKB |
| FOXL2 | CBLN4 | SPARC | CST1 | PSMB8 |
| GAS5 | NAT10 | SATB1 | THBS2 | ADAMTS9 |
| YY1 | ESRP1 | LMNB2 | ACRBP | C15orf52 |
| HP | MTHFD1 | TDG | FBXL2 | EPHB3 |
| GZMB | PVR | SLC35F5 | FFAR4 | B3GNT8 |

Figura 7: Método de extracción para aislar biomarcadores de ARN eucariota en muestras de heces. A. trazas de BioAnalyzer para cuatro muestras de extracción de ARN felino y cuatro muestras de extracción de ARN canino. B. números de integridad de ARN (RIN) para todas las ocho muestras de extracción de ARN. C. concentración de ARN eucariota total estimado para todas las ocho muestras de extracción de ARN.

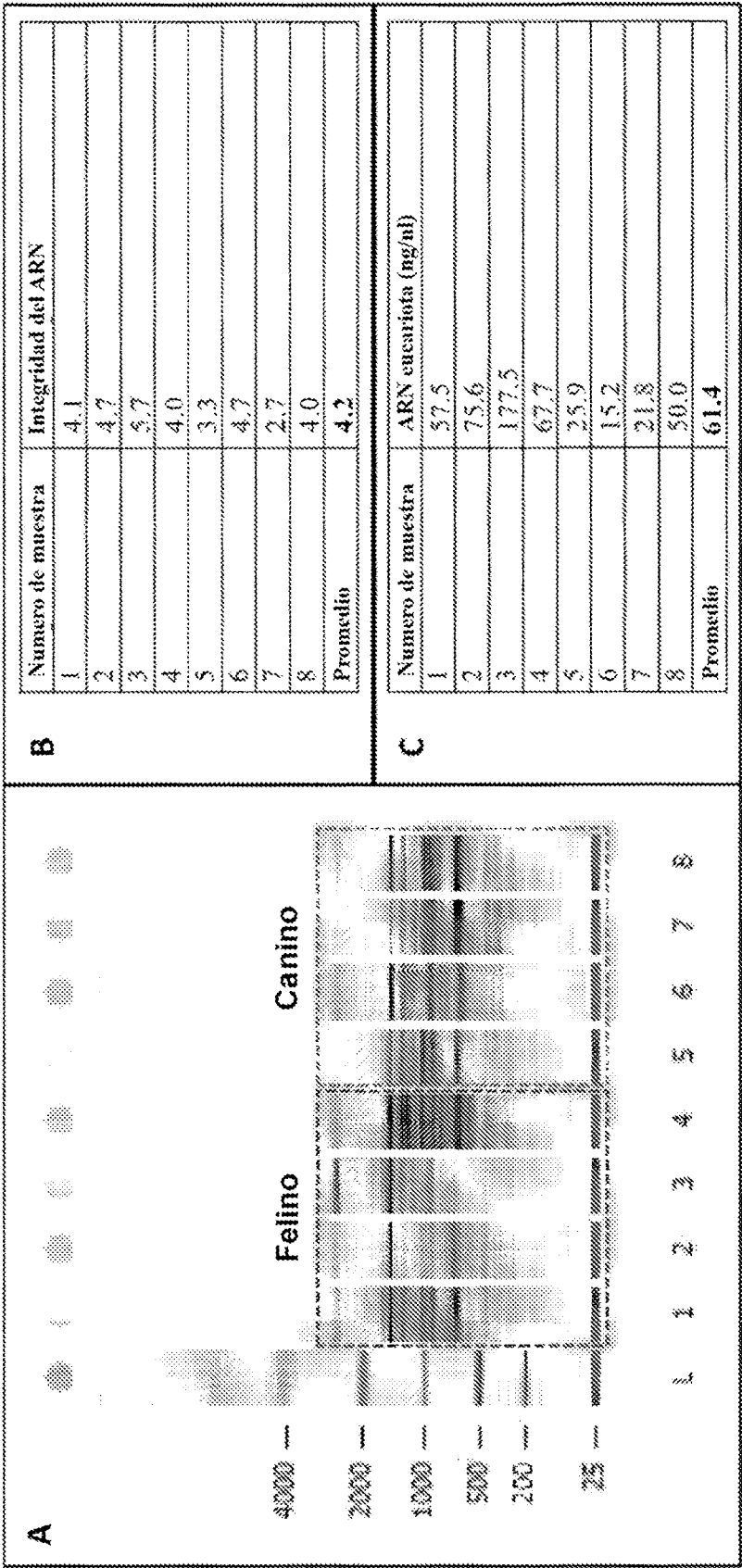


Figura 8: Resultados de RT-qPCR de muestras caninas A. Resultados de RT-qPCR para muestras positivas de C-mu en cuatro muestras caninas. B. Resultados de RT-qPCR para el ARN específico de reordenamientos de células T en cuatro muestras caninas.

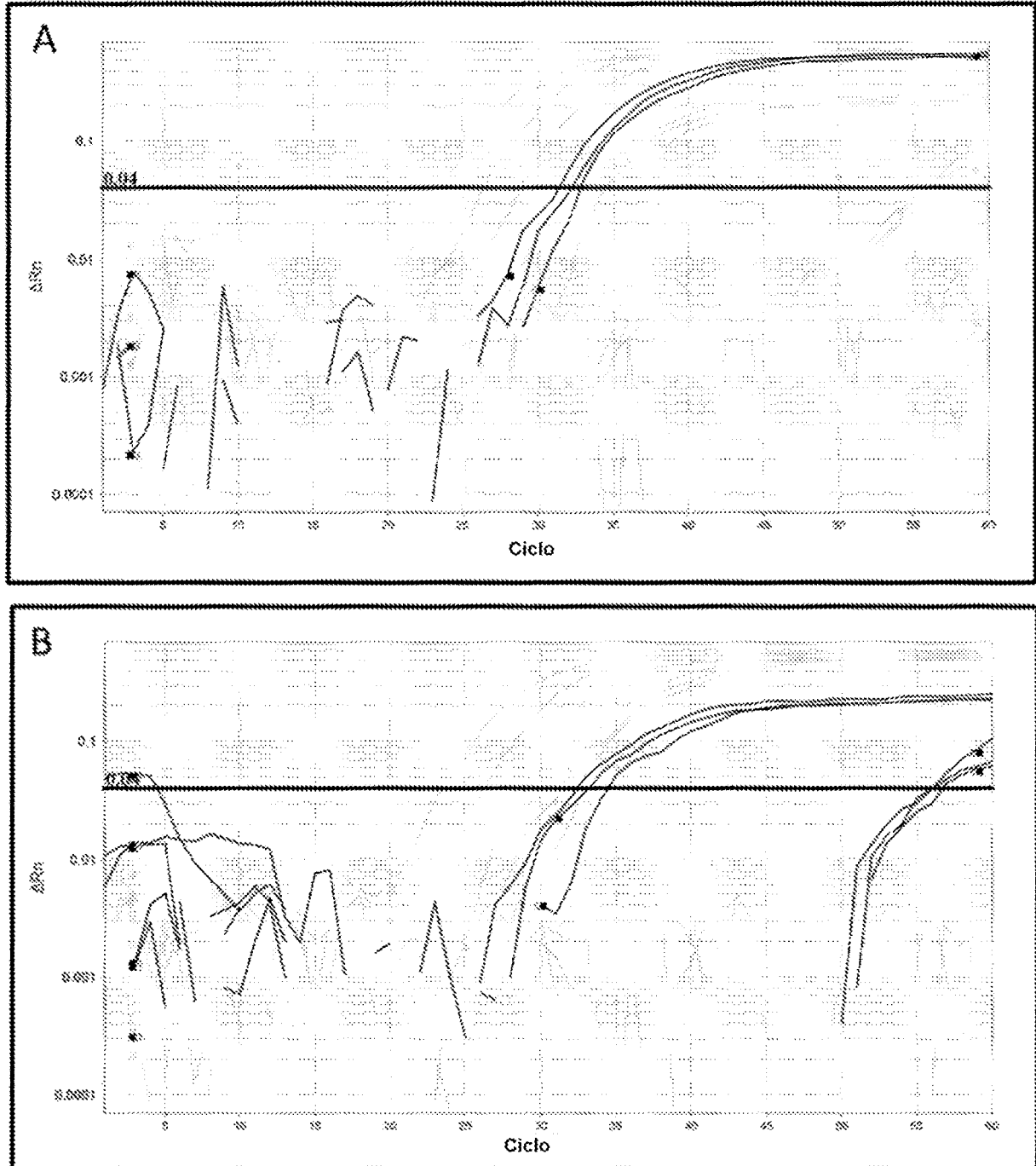


Figura 9: Análisis de RT-qPCR de muestra felina para la actina-B de control positivo en cuatro muestras felinas

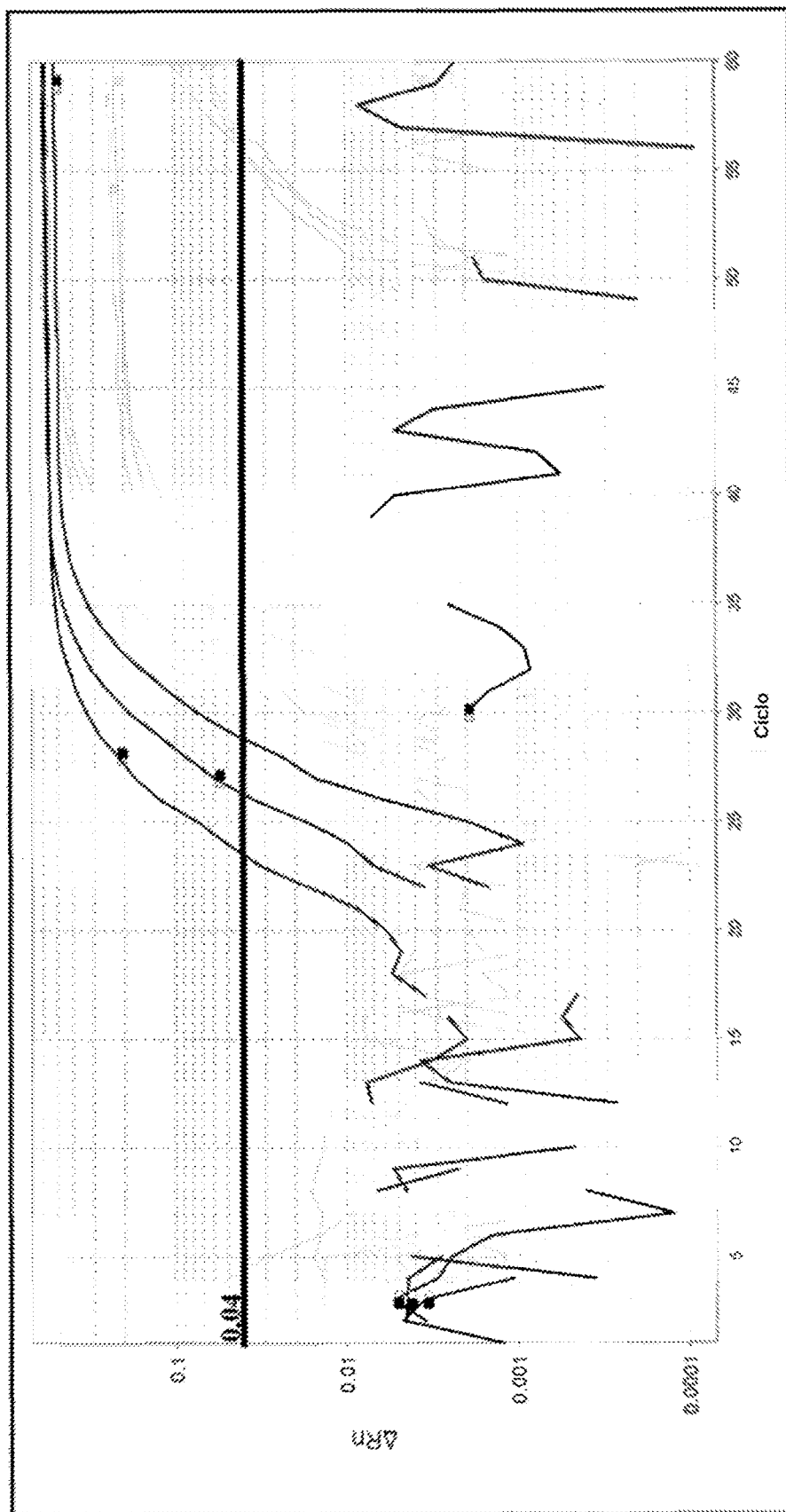


Figura 10: Análisis de replicación para todos los 70,524 grupos de transcripción A) correlaciones de cuatro réplicas técnicas. B) correlaciones de cuatro replicas biológicas. C) correlaciones de replicas técnicas separadas por tiempo.

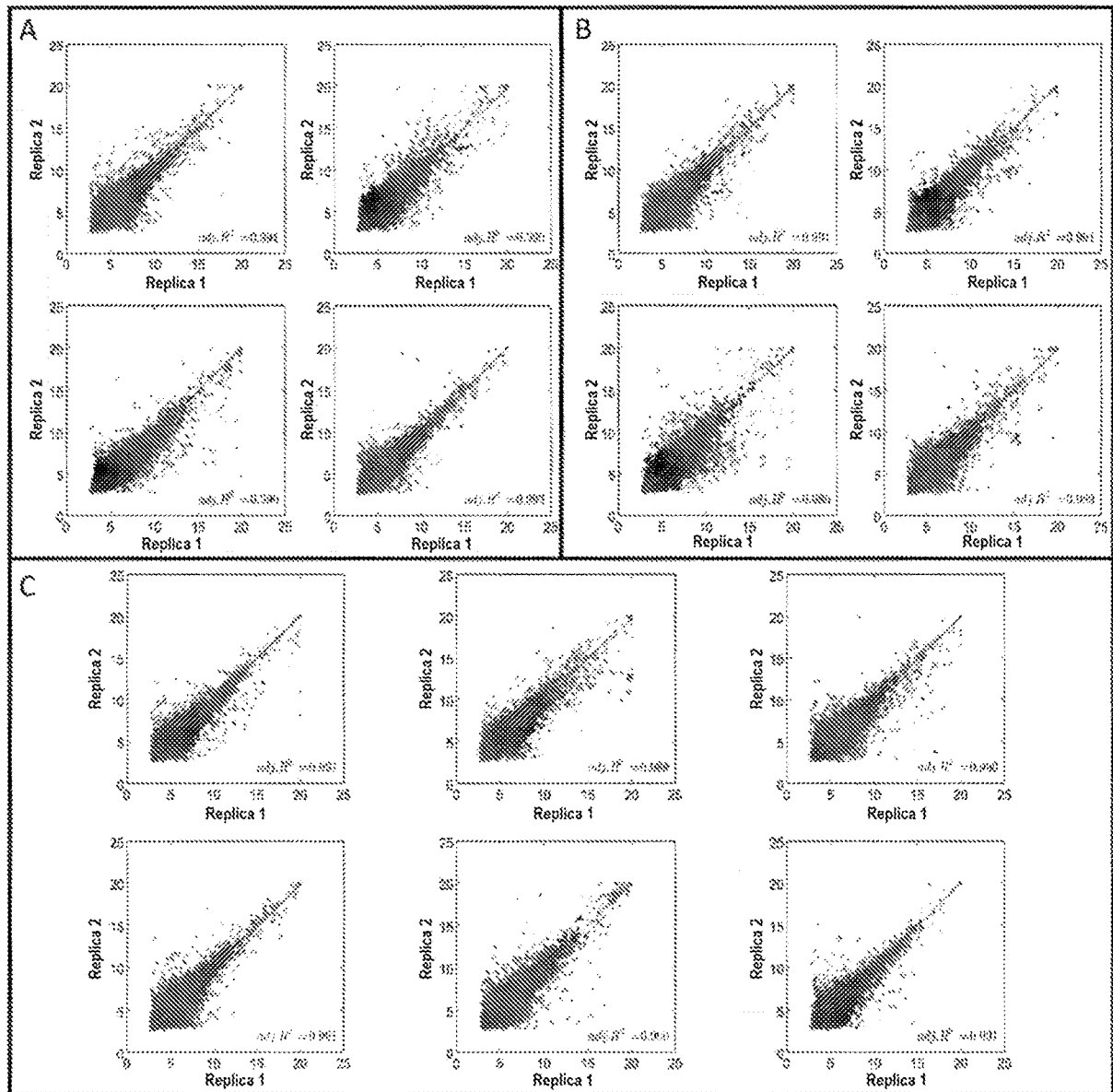


Figura 11: Análisis de replicación para 5,149 agrupaciones de transcripción utilizando búsqueda de bibliografía A) correlaciones de cuatro replicas técnicas. B) correlaciones de cuatro réplicas biológicas. C) correlaciones de réplicas técnicas separadas por tiempo. De izquierda superior a derecha inferior: muestra #70 saludable, muestra #92 adenoma, muestra #98 adenoma, muestra #105 adenoma, muestra de cáncer colorrectal estadio II #34, muestra de cáncer colorrectal estadio IV #36.

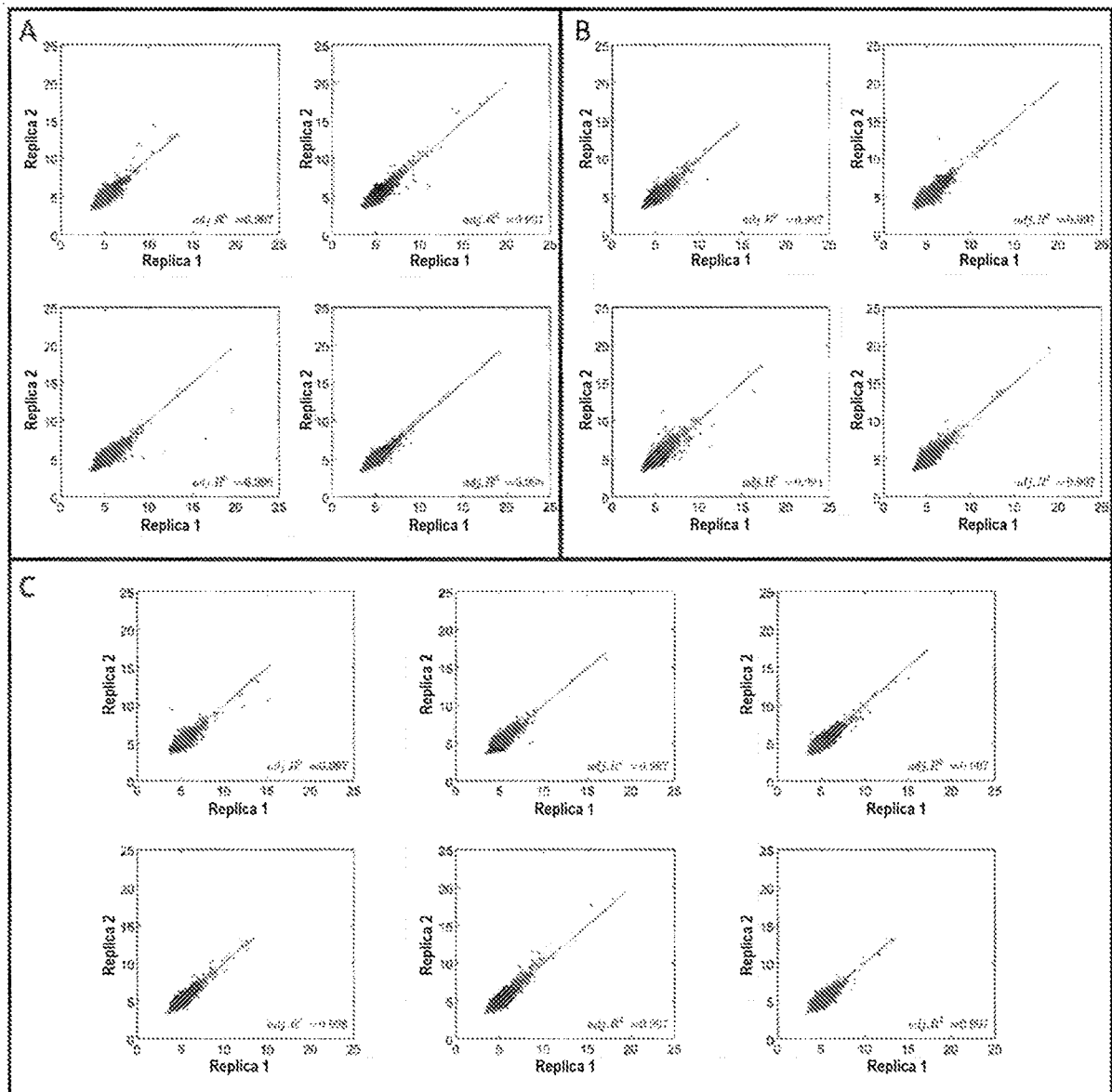


Figura 12: método de extracción para aislar biomarcadores de ARNm humano en muestras de heces. A. trazas de BioAnalyzer para los métodos de extracción de ácido nucleico descritos en la bibliografía B. trazas de BioAnalyzer para el método de extracción de ácido nucleico descrito en la invención.

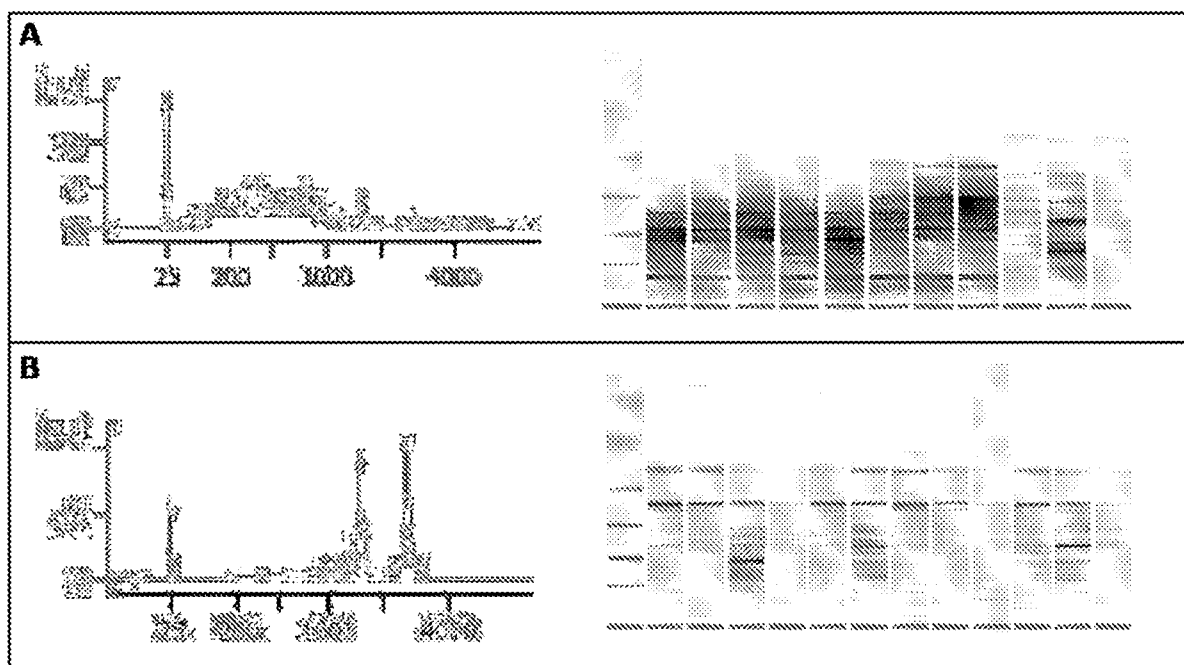
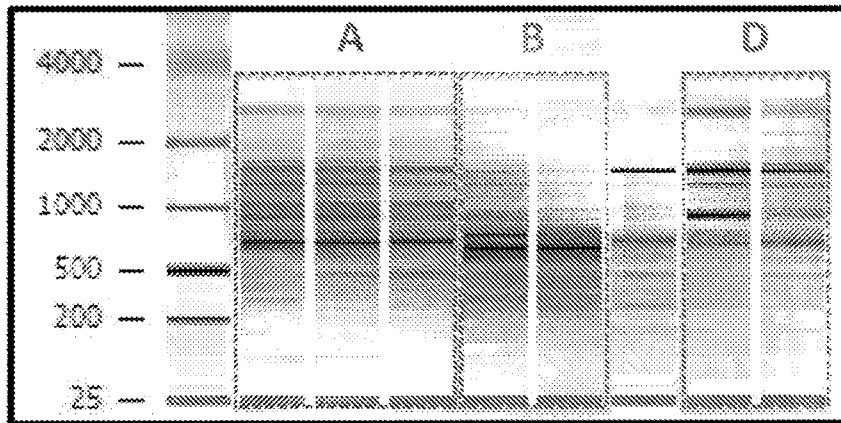


Figura 13: 400 firmas de gen para cáncer colorrectal y diagnóstico de pólipos precancerígenos

| | | | | | | | |
|----------|---------|---------|-----------|---------|----------|----------|-----------|
| ABCB1 | CD151 | DDIAS | GAST | KIF2C | NOS2 | RALGDS | TCF20 |
| ABCG2 | CD24 | DDIT3 | GDNF | KISS1R | NUP88 | RARB | TCF7 |
| ACHE | PDL1 | DHRS9 | GLI2 | KLF4 | OAS1 | RASAL2 | TCF7L2 |
| ACRBP | CD276 | DPP6 | GLRX3 | KLK4 | ODC1 | RASSF1 | TCN2 |
| ACTB | CD3D | DPYD | GNB2 | KMT2C | OGG1 | RBM47 | TDG |
| ACY1 | CD8A | DRD4 | GNG12 | KRAS | OLR1 | RFC3 | TENM1 |
| ADAMTS9 | CDC20 | DSC2 | GNG8 | KRT19 | OTX1 | RGCC | TERT |
| ADCY1 | CDC42 | DST | GOLPH3 | KRT20 | PARP1 | RHOA | TGFA |
| ADCY5 | CDC73 | EDN1 | GSN | LAMC2 | PARP11 | RNU2-1 | THBD |
| AGR2 | CDH1 | EGFR | GSTA4 | LAMP2 | PDCD11G2 | ROCK2 | TIAM1 |
| AHNAK | CDH22 | EGLN2 | GSTM1 | LBR | PDGFA | ROR1 | TJP3 |
| AIM2 | CDHR5 | EGLN3 | GSTP1 | LDHA | PDGFRB | RPIA | TLR2 |
| AKT1 | CDK1 | EHMT2 | HDAC1 | LGR4 | PEA3 | RPL19 | TLR3 |
| AKT2 | CDK12 | ELAVL1 | HDAC11 | LGR5 | PER3 | RPLP0 | TMEFF2 |
| ALDOA | CDKN1A | ELF3 | HDAC5 | LIN28A | PGGT1B | RPRM | TMEM158 |
| ALK | CDKN2A | EMP1 | HDHD3 | LMNB2 | PIK3CB | RPS27A | TMPPSS2 |
| AMD1 | CDKN2B | EPCAM | HGF | LPCAT1 | PLA2G2A | RRAS2 | TNFR |
| ANPEP | CEACAM1 | EPHA2 | HIF1A | LRP5 | PLAUR | RRM2 | TNFRSF10B |
| ANXA2P2 | CEACAM5 | EPHA7 | HIST1H2BM | LRPAP1 | PLCD1 | RTN4 | TNFRSF1A |
| APC | CEACAM6 | EPHB2 | HK1 | LTA | PLCG1 | S100A14 | TNFSF10 |
| APTX | CGN | EPOR | HK2 | LY96 | PLIN2 | SALL3 | TNFSF9 |
| AREG | CHD7 | EPSBL1 | HLA-DQB1 | MAGED4 | PLS1 | SALL4 | TNS4 |
| ARL4C | CHP1 | ERBB3 | HLA-DRB1 | MAP2K3 | PMS2 | SATB1 | TOP1 |
| ASCL1 | CHRNA7 | ERBB2IP | HMGA1 | MAPK3 | POLD1 | SCGB2A1 | TP73 |
| ATM | CIB1 | ERCC1 | HMGA2 | MDK | POLE | SCNN1A | TSPAN1 |
| ATP11A | CISH | ESR1 | HMGCR | MET | POU2F1 | SCNN1B | TSPEAR |
| ATP1B1 | CITED1 | ESR2 | HMOX1 | METAP2 | PPARGC1A | SEMA4A | TTL3 |
| ATPIF1 | CLDN3 | ESRP1 | HXA10 | MKI67 | PPFIBP2 | SEMG1 | TUBB3 |
| ATRX | CLDN4 | FASN | HXA9 | MKRN3 | PPM1D | SETD2 | TXNIP |
| AURKA | CLDN7 | FBXL2 | HP | MLH1 | PPP2CB | SFN | TYMP |
| B2M | CNTN1 | FBXL20 | HRH2 | MLH3 | PPP2R1B | SFRP5 | TYMS |
| B3GNT3 | COL5A1 | FCGR3A | HSD17B1 | MXI1 | PRICKLE2 | SHROOM2 | UBA52 |
| B4GALT3 | COMT | FCN2 | HSPB6 | MPO | PRKAR1A | SKP1 | UBD |
| BAD | CREB5 | FDFT1 | ID1 | MSH5 | PROK2 | SLC12A5 | UGCG |
| BCR | CREBBP | FDPS | IFNA1 | MT-CO1 | PRSS22 | SLC25A46 | UGT2B7 |
| BIRC5 | CSF1R | FFAR4 | IFNAR1 | MTA1 | PSMB8 | SLC35F5 | UHRF1 |
| BMP3 | CSF3R | FGF4 | IGF2 | MTHFD1 | PSMD10 | SMAD7 | VEGFA |
| BRAF | CSRNP3 | FHIT | IGF2R | MUC1 | PTEN | SNX1 | VEZF1 |
| BRS3 | CST1 | FLCN | IGFBP3 | MUC2 | PTGER2 | SPATA2 | VTCN1 |
| BUB1 | CTLA4 | FNTA | IL1B | MUC6 | PTGES2 | SPINT1 | WDR49 |
| BVES | CTNNA1 | FNTB | IL1RN | MYNN | PTGS2 | SPOP | WNT3A |
| C21orf33 | CTNNA1 | FOS | IL2 | MYO18B | PTK2B | SRC | WNT6 |
| CA6 | CTNND1 | FOXA2 | IL23A | NAPSA | PTK7 | SRGN | WNT7A |
| CABLES1 | CUX1 | FOXH1 | IRF6 | NAT10 | PTPN11 | SSTR2 | WNT9A |
| CAMK2D | CXCL8 | FOXO4 | ITCH | NAT2 | PTPRS | ST14 | WRN |
| CASC2 | CXCR2 | FRMPD2 | ITGA3 | NCK1 | PTPRU | STAB1 | YY1 |
| CBLN4 | CXCR4 | FUT3 | ITGB4 | NDOR1 | PVR | STAT3 | ZNF135 |
| CCDC129 | CYC5 | GADD45A | JAK1 | NEUROG1 | PYCR1 | STRK1 | ZNF492 |
| CCNB1 | CYP1A1 | GAPDH | KALRN | NFIB | PYGO1 | SUZ12 | ZNF560 |
| CCNG2 | DBP | GAS5 | KIF26B | NFKBIA | RAC1 | TAC1 | ZSCAN15 |

Figura 14: Trazas de BioAnalyzer para réplicas biológicas para cuatro diferentes muestras de felino. Las columnas uno a la tres son de gato "A", las columnas cuatro y cinco son de gato "B", la columna seis es del gato "C", y la columna siete y ocho son del gato "D".



| Figura 15B: Tablas que representan los cebadores designados y utilizados para RT-qPCR de muestras de felino | | | | |
|--|-------------------------------|-------------------------------|--------------------|---|
| B | Nombre del producto | Diana | Nombre del cebador | Secuencia (5-3') SEQ ID NO: |
| ActB | Control de célula epitelial | Control de célula epitelial | ActBfwd | GCCAACCGTGAGAAGATGAC |
| | | | ActBrev | GGAGAAGAACAGAACCCCTGG |
| GAPDH | Control de normalización | Control de normalización | GDHfwd | CGGATTGGCCGTATTGGGCGC |
| | | | GDHrev | GGTGAATCATACTGGAACATGTAGACCA |
| Región gamma del receptor de célula T | Cebadores de la región V | Cebadores de la región V | fcrgv1-2 | GSAGAAAGAGCGACGAGGGCGTG |
| | | | fcrgv3 | GGCGAAAGAGCGATGAGGGAGTG |
| | | | fcrgv4 | GTAGTGAGGAGRATGCTGGTCTG |
| | | | fcrgvS | GGCAGAAAGCATGACAAAGGGCATG |
| | | | TCR felino unirev | AGKTTTMYTTCAGCAATTGAAGG |
| | | | fcrgvj1 | [FAM] CCCTGAGCAGTGTGCCAGSAC [BHQ1] |
| Región 3 determinante complementaria de célula B | Sondas TaqMan de la región J | Sondas TaqMan de la región J | fcrvj2 | [CY3] GGGGGAGTTACKATGASCTTARTTCC [BHQ1] |
| | | | fcrvj3 | [JOE] ATCCAGATCTCAGGTTTGGAGGAGG [BHQ1] |
| | | | fCDR3f1 | CCAGGCTCCAGGGAAGGG |
| | | | fCDR3f2 | TCCAGAGACAACGCCAGGAAC |
| | Cebadores de la región VH | Cebadores de la región VH | fel CDR3 unirev | GTTCGRCRCRTRCTSGTTTC |
| | | | fCDR3jh1 | [FAM] ACACCGTCACCAGGGCTCC [BHQ1] |
| | Sondas TaqMan de la región JH | Sondas TaqMan de la región JH | fCDR3jh2 | [CY3] TGAGGACACTGTGACTAATGGTTCC [BHQ1] |
| | | | fCDR3jh3 | [JOE] GGACACCGTCACYAKGVYTCC [BHQ1] |
| | | | | |
| | | | | |

Figura 16: Amplificación PCR de transcritos relacionados con linfocito. A) resultados de RT-qPCR para ARN específico de reordenamiento de célula B en cuatro muestras de canino. B) resultados de RT-qPCR para ARN específico de reordenamiento de célula T en cuatro muestras de felino.

