

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 976 360**

51 Int. Cl.:

C12Q 1/6886 (2008.01)

C40B 40/06 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **27.01.2017 PCT/AU2017/050066**

87 Fecha y número de publicación internacional: **03.08.2017 WO17127893**

96 Fecha de presentación y número de la solicitud europea: **27.01.2017 E 17743495 (8)**

97 Fecha y número de publicación de la concesión europea: **20.03.2024 EP 3408412**

54 Título: **Métodos para evaluar el riesgo de desarrollar cáncer colorrectal**

30 Prioridad:

28.01.2016 AU 2016900254
16.08.2016 AU 2016903246

45 Fecha de publicación y mención en BOPI de la
traducción de la patente:
30.07.2024

73 Titular/es:

THE UNIVERSITY OF MELBOURNE (100.0%)
Royal Parade
Parkville, Victoria 3010, AU

72 Inventor/es:

JENKINS, MARK;
BUCHANAN, DANIEL y
HOPPER, JOHN L.

74 Agente/Representante:

IZQUIERDO BLANCO, María Alicia

Observaciones:

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

ES 2 976 360 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

Métodos para evaluar el riesgo de desarrollar cáncer colorrectal

5 **CAMPO DE LA INVENCION**

[0001] La presente divulgación se refiere a métodos y sistemas para evaluar el riesgo de un sujeto humano de desarrollar cáncer colorrectal. Estos métodos pueden combinarse con el riesgo clínico de los sujetos para mejorar el análisis del riesgo. Estos métodos pueden utilizarse para ayudar a tomar decisiones sobre los regímenes de cribado del cáncer colorrectal adecuados.

10 **FONDO DE LA INVENCION**

[0002] Los programas de cribado del cáncer colorrectal propugnan la administración de pruebas a individuos de poblaciones aparentemente sanas para identificar a los individuos que padecen cáncer colorrectal premaligno o en estadios tempranos, de modo que puedan beneficiarse de la prevención o de un tratamiento precoz. Las pruebas de cribado pueden incluir pruebas de sangre oculta en heces y colonoscopia. En la población de riesgo medio, el cribado basado en pruebas de sangre oculta en heces reduce la mortalidad colorrectal entre un 15% y un 25% (Hewitson et al., 2007). El cribado endoscópico puede reducir la mortalidad entre un 30% y un 40% (Brenner et al., 2014).

[0003] Examinar a un gran número de personas puede ser costoso. Lo ideal sería que la decisión sobre quién debe someterse al cribado, así como el procedimiento y la intensidad del mismo, se basaran en el riesgo de cáncer colorrectal de cada persona. Sin embargo, dado que actualmente no existen métodos precisos o válidos para determinar el riesgo individual de padecer la enfermedad, el cribado selectivo sólo se basa en los factores de riesgo muy generales de la edad, el sexo y, en ocasiones, los antecedentes familiares. Esto hace que los programas de cribado sean ineficaces porque muchos de los que se someten a cribado nunca padecerán cáncer colorrectal, y muchos de los que no se someten a cribado corren un riesgo considerable de padecer la enfermedad (Ait Ouakrim et al., 2012).

[0004] Las evaluaciones del riesgo genético pueden aumentar la eficacia de los programas de cribado. Sin embargo, la susceptibilidad genética al cáncer colorrectal hereditario es compleja e implica múltiples variantes y genes.

[0005] Para aumentar la eficacia del cribado y reducir la mortalidad por cáncer colorrectal, se necesitan métodos mejorados para evaluar el riesgo de que un sujeto humano desarrolle cáncer colorrectal.

[0006] HSU et al. ("A model to determine colorectal cancer risk using common genetic susceptibility loci", GASTROENTEROLOGY, vol. 148, nº 7, 2015, páginas 1330-1339) divulga un método para evaluar el riesgo de un sujeto humano de desarrollar cáncer colorrectal. El método comprende el genotipado de un conjunto de 27 SNP seleccionados de entre las variantes de riesgo previamente comunicadas y validadas (véase la Tabla Suplementaria 4) en una muestra biológica del sujeto, la determinación de la presencia de variantes de riesgo y el cálculo del riesgo o predisposición al cáncer colorrectal (véase el Resumen y Métodos).

40 **RESUMEN DE LA INVENCION**

[0007] Los presentes inventores han identificado SNP dentro del genoma que son útiles para evaluar el riesgo de que un sujeto desarrolle cáncer colorrectal.

[0008] En consecuencia, la presente invención se refiere a un método para evaluar el riesgo de un sujeto humano de desarrollar cáncer colorrectal que comprende realizar una evaluación del riesgo genético del sujeto, en la que la evaluación del riesgo genético implica detectar, en una muestra biológica derivada del sujeto, la presencia de al menos 45 polimorfismos de nucleótido único proporcionados en la Tabla 1. La presente invención se define además en las reivindicaciones.

[0009] También se divulga en el presente documento a un método para evaluar el riesgo de un sujeto humano de desarrollar cáncer colorrectal que comprende: realizar una evaluación del riesgo genético del sujeto, donde la evaluación del riesgo genético implica detectar, en una muestra biológica derivada del sujeto, la presencia de al menos 28 polimorfismos de nucleótido único seleccionados de la Tabla 1, o un polimorfismo de nucleótido único en desequilibrio de ligamiento con uno o más de ellos.

[0010] Algunos polimorfismos de nucleótido único son más informativos que otros para una evaluación de riesgo concreta. Así, la evaluación del riesgo genético comprende al menos detectar la presencia de los polimorfismos de nucleótido único rs3987, rs35509282 y rs744166. Como se describe en el presente documento, puede detectarse la presencia de un polimorfismo de nucleótido único en desequilibrio de ligamiento con uno o más de los polimorfismos de nucleótido único rs3987, rs35509282 y rs744166.

[0011] La evaluación del riesgo genético puede comprender la detección de más de 28 polimorfismos de nucleótido único seleccionados de la Tabla 1, o un polimorfismo de nucleótido único en desequilibrio de ligamiento con uno o más de ellos.

Por ejemplo, pueden detectarse al menos 29, al menos 30, al menos 31, al menos 32, al menos 33, al menos 34, al menos 35, al menos 36, al menos 37, al menos 38, al menos 39, al menos 40, al menos 41, al menos 42, al menos 43, al menos 44 polimorfismos de nucleótido único. En el método de la presente invención, se detectan al menos 45 polimorfismos de nucleótido único.

[0012] La evaluación del riesgo genético comprende la detección de la presencia del polimorfismo de nucleótido único rs5934683. Como se describe en el presente documento, puede detectarse la presencia de un polimorfismo de nucleótido único en desequilibrio de ligamiento con el polimorfismo de nucleótido único rs5934683.

[0013] En el método de la presente invención, la evaluación del riesgo genético puede combinarse con una evaluación del riesgo clínico para obtener el riesgo de un sujeto humano de desarrollar cáncer colorrectal. En un ejemplo, la evaluación del riesgo clínico implica obtener información del sujeto sobre uno o más de los siguientes aspectos: antecedentes médicos de cáncer colorrectal, edad, antecedentes familiares de cáncer colorrectal, resultados de una colonoscopia o sigmoidoscopia previa y raza/etnia. En otro ejemplo, la evaluación del riesgo clínico implica obtener información del sujeto sobre la edad y/o los antecedentes de cáncer colorrectal de un familiar de primer grado. En una realización, los antecedentes familiares de cáncer colorrectal incluyen antecedentes familiares multigeneracionales.

[0014] Un experto en la materia apreciará que la combinación de la evaluación del riesgo clínico y la evaluación del riesgo genético define el riesgo general de los sujetos de desarrollar cáncer de colon. Así, los métodos de la invención pueden utilizarse para evaluar el riesgo global.

[0015] Los métodos aquí descritos pueden determinar el riesgo absoluto de un sujeto humano femenino de desarrollar cáncer de colon.

[0016] Alternativamente, los métodos aquí descritos pueden determinar el riesgo relativo de un sujeto humano femenino de desarrollar cáncer de colon.

[0017] El método de la invención puede ser aplicable a sujetos con síntomas de cáncer colorrectal. Por ejemplo, los sujetos que han tenido una prueba de sangre oculta en heces positiva pueden ser evaluados utilizando los métodos de la presente divulgación. Las pruebas de sangre oculta en heces se recomiendan generalmente a los sujetos de unos 50 años de edad. Los presentes inventores han descubierto que ciertos individuos tienen un mayor riesgo de padecer cáncer colorrectal mucho antes de alcanzar los 50 años de edad, en particular si a un pariente de primer grado se le ha diagnosticado cáncer colorrectal. Estos resultados sugieren que algunos individuos deberían ser evaluados antes para determinar si presentan riesgo de cáncer colorrectal. Así, los sujetos evaluados mediante los métodos aquí descritos pueden tener al menos 40 años de edad. En algunas realizaciones, el sujeto evaluado tiene al menos 30 años de edad y tiene antecedentes familiares de cáncer colorrectal. Por ejemplo, un familiar de primer grado puede haber sido diagnosticado de cáncer colorrectal.

[0018] En los métodos de la invención, el sujeto puede ser hombre o mujer. En algunas realizaciones, el sujeto es varón.

[0019] Los sujetos que se determina que están en riesgo de desarrollar cáncer colorrectal utilizando la presente invención pueden entonces inscribirse en un programa de cribado o someterse a un cribado más frecuente.

[0020] En algunas realizaciones, el rendimiento del método de la invención se caracteriza por un área bajo la curva (AUC) de al menos 0,63 aproximadamente.

[0021] Un polimorfismo de nucleótido único en desequilibrio de ligamiento puede tener un desequilibrio de ligamiento superior a 0,9. Alternativamente, un polimorfismo de un solo nucleótido en desequilibrio de ligamiento puede tener un desequilibrio de ligamiento de 1.

[0022] En otro aspecto, los métodos de la presente invención se utilizan para determinar la necesidad de pruebas diagnósticas rutinarias de un sujeto humano para cáncer colorrectal. Por ejemplo, si se tiene en cuenta que cada uno de los polimorfismos de nucleótido único puede estar presente hasta dos veces en el genoma diploide somático del sujeto, un sujeto que tenga al menos 41, al menos 42, al menos 44, al menos 46, al menos 50, al menos 55, al menos 60, al menos 65, o al menos 70, de los polimorfismos de nucleótido único debería inscribirse en un programa de cribado fecal oculto, colonoscópico o sigmoidoscópico. Alternativamente, si la evaluación sitúa al sujeto en el 20% superior de sujetos de una población con riesgo de desarrollar cáncer colorrectal, el sujeto se inscribe en un programa de cribado oculto en heces, colonoscópico o sigmoidoscópico. Como alternativa adicional, si la evaluación sitúa al sujeto en el 10% superior de sujetos de una población con riesgo de desarrollar cáncer colorrectal, el sujeto se inscribe en un programa de cribado oculto en heces, colonoscópico o sigmoidoscópico.

[0023] En otro aspecto, la presente invención proporciona un método de cribado de cáncer colorrectal en un sujeto humano, el método comprende evaluar el riesgo del sujeto de desarrollar cáncer colorrectal utilizando el método de la invención, y cribado rutinario de cáncer colorrectal en el sujeto si se evalúa que tiene riesgo de desarrollar cáncer colorrectal.

[0024] En otro aspecto, los métodos de la presente divulgación se utilizan como una terapia contra el cáncer colorrectal para su uso en la prevención del cáncer colorrectal en un sujeto humano en riesgo del mismo.

5 [0025] También se divulga en el presente documento, pero no forma parte de la invención reivindicada, un kit que comprende al menos 28 conjuntos de cebadores para amplificar 28 o más ácidos nucleicos, en el que los 28 o más ácidos nucleicos comprenden un polimorfismo de nucleótido único seleccionado de la Tabla 1, o un polimorfismo de nucleótido único en desequilibrio de ligamiento con uno o más de ellos.

10 [0026] También se divulga en el presente documento, pero no forma parte de la invención reivindicada, una matriz genética que comprende al menos 28 conjuntos de sondas para hibridar con 28 o más ácidos nucleicos, en los que los 28 o más ácidos nucleicos comprenden un polimorfismo de nucleótido único seleccionado de la Tabla 1, o un polimorfismo de nucleótido único en desequilibrio de ligamiento con uno o más de ellos.

15 [0027] También se divulga en el presente documento un método implementado por ordenador para evaluar el riesgo de un sujeto humano de desarrollar cáncer colorrectal, el método operable en un sistema informático que comprende un procesador y una memoria, el método comprende:

20 recibir datos de riesgo genético del sujeto, en los que los datos de riesgo genético se obtuvieron detectando, en una muestra biológica derivada del sujeto, la presencia de al menos 28 polimorfismos de nucleótido único de la Tabla 1, o un polimorfismo de nucleótido único en desequilibrio de ligamiento con uno o más de ellos;
procesar los datos para determinar el riesgo del sujeto humano de desarrollar cáncer colorrectal;
emitir el riesgo del sujeto humano de desarrollar cáncer colorrectal.

25 [0028] En una realización, el método implementado por ordenador comprende además la recepción de datos de riesgo clínico del sujeto;

30 procesar los datos para combinar los datos de riesgo clínico con los datos de riesgo genético para obtener el riesgo del sujeto de desarrollar cáncer colorrectal;
emitir el riesgo del sujeto de desarrollar cáncer colorrectal.

En otro aspecto de la invención, se proporciona un método implementado por ordenador para evaluar el riesgo de un sujeto humano de desarrollar cáncer colorrectal, el método operable en un sistema informático que comprende un procesador y una memoria, el método comprende:

35 recibir datos de riesgo genético del sujeto, en los que los datos de riesgo genético se obtuvieron detectando, en una muestra biológica derivada del sujeto, la presencia de al menos 45 polimorfismos de nucleótido único de la Tabla 1, o un polimorfismo de nucleótido único en desequilibrio de ligamiento con uno o más de ellos;
40 procesar los datos para determinar el riesgo del sujeto humano de desarrollar cáncer colorrectal;
emitir el riesgo del sujeto humano de desarrollar cáncer colorrectal.

[0029] En una realización, el método implementado por ordenador comprende además la recepción de datos de riesgo clínico del sujeto;

45 procesar los datos para combinar los datos de riesgo clínico con los datos de riesgo genético para obtener el riesgo del sujeto de desarrollar cáncer colorrectal;
emitir el riesgo del sujeto de desarrollar cáncer colorrectal.

50 [0030] En algunas realizaciones, los datos de riesgo del sujeto se reciben desde una interfaz de usuario acoplada al sistema informático. En algunas realizaciones, los datos de riesgo del sujeto se reciben de un dispositivo remoto a través de una red de comunicaciones inalámbrica. En algunas realizaciones, la interfaz de usuario o el dispositivo remoto es una plataforma de matriz SNP. En algunas realizaciones, la salida comprende la salida de información a una interfaz de usuario acoplada al sistema informático. En algunas realizaciones, la salida comprende la transmisión de información a un dispositivo remoto a través de una red de comunicaciones inalámbrica.

55 [0031] Cualquier ejemplo aquí expuesto se considerará aplicable *mutatis mutandis* a cualquier otro ejemplo, a menos que se indique específicamente lo contrario.

60 [0032] La presente divulgación de la invención no debe limitarse en su alcance por los ejemplos específicos descritos en el presente documento, que están destinados únicamente a fines de ejemplificación. Los productos, composiciones y métodos funcionalmente equivalentes están claramente dentro del alcance de la divulgación, tal y como se describe en el presente documento.

65 [0033] A lo largo de esta especificación, a menos que se indique específicamente lo contrario o que el contexto requiera otra cosa, la referencia a una única etapa, composición de materia, grupo de etapas o grupo de composiciones de materia

se entenderá que abarca uno y una pluralidad (es decir, uno o más) de dichas etapas, composiciones de materia, grupos de etapas o grupo de composiciones de materia.

[0034] A lo largo de esta especificación, se entenderá que la palabra "comprende", o variaciones como "comprende" o "que comprende", implica la inclusión de un elemento, número entero o etapa, o grupo de elementos, números enteros o etapas, pero no la exclusión de cualquier otro elemento, número entero o etapa, o grupo de elementos, números enteros o etapas.

[0035] La invención se describe a continuación mediante los siguientes Ejemplos no limitativos y con referencia a los dibujos adjuntos.

BREVE DESCRIPCIÓN DE LOS DIBUJOS ADJUNTOS

[0036]

Figura 1. La distribución simulada de alelos de riesgo para 1.000.000 de personas con antecedentes de cáncer colorrectal (rojo) y 1.000.000 de personas sin antecedentes de cáncer colorrectal (azul); y el riesgo acumulado de cáncer colorrectal hasta los 70 años para el número de alelos de riesgo para una población australiana (cuadrado) y estadounidense (círculo).

Figura 2. Riesgos australianos de cáncer colorrectal (hombres y mujeres combinados) por categoría de edad, antecedentes familiares de cáncer colorrectal (pariente de primer grado) y por número de alelos de riesgo. Panel A: riesgos acumulados hasta los 70 años con los quintiles más altos y más bajos para el número de alelos de riesgo. Panel B: riesgos acumulados hasta los 70 años con deciles más altos y más bajos para el número de alelos de riesgo. Panel C: Riesgos a 5 años con los quintiles más altos y más bajos para el número de alelos de riesgo. Panel D: Riesgos a 5 años con deciles más altos y más bajos para el número de alelos de riesgo.

Figura 3. Riesgos de cáncer colorrectal en EE.UU. (hombres y mujeres combinados) por categoría de edad, antecedentes familiares de cáncer colorrectal (pariente de primer grado) y por número de alelos de riesgo. Panel A: riesgos acumulados hasta los 70 años con los quintiles más altos y más bajos para el número de alelos de riesgo. Panel B: riesgos acumulados hasta los 70 años con deciles más altos y más bajos para el número de alelos de riesgo. Panel C: Riesgos a 5 años con los quintiles más altos y más bajos para el número de alelos de riesgo. Panel D: Riesgos a 5 años con deciles más altos y más bajos para el número de alelos de riesgo.

Figura 4. Riesgos australianos de cáncer colorrectal (varones) por categoría de edad, antecedentes familiares de cáncer colorrectal (pariente de primer grado) y por número de alelos de riesgo. Panel A: riesgos acumulados hasta los 70 años con los quintiles más altos y más bajos para el número de alelos de riesgo. Panel B: riesgos acumulados hasta los 70 años con deciles más altos y más bajos para el número de alelos de riesgo. Panel C: Riesgos a 5 años con los quintiles más altos y más bajos para el número de alelos de riesgo. Panel D: Riesgos a 5 años con deciles más altos y más bajos para el número de alelos de riesgo.

Figura 5. Riesgos australianos de cáncer colorrectal (mujeres) por categoría de edad, antecedentes familiares de cáncer colorrectal (pariente de primer grado) y por número de alelos de riesgo. Panel A: riesgos acumulados hasta los 70 años con los quintiles más altos y más bajos para el número de alelos de riesgo. Panel B: riesgos acumulados hasta los 70 años con deciles más altos y más bajos para el número de alelos de riesgo. Panel C: Riesgos a 5 años con los quintiles más altos y más bajos para el número de alelos de riesgo. Panel D: Riesgos a 5 años con deciles más altos y más bajos para el número de alelos de riesgo.

Figura 6. Riesgos de cáncer colorrectal en EE.UU. (varones) por categoría de edad, antecedentes familiares de cáncer colorrectal (pariente de primer grado) y por número de alelos de riesgo. Panel A: riesgos acumulados hasta los 70 años con los quintiles más altos y más bajos para el número de alelos de riesgo. Panel B: riesgos acumulados hasta los 70 años con deciles más altos y más bajos para el número de alelos de riesgo. Panel C: Riesgos a 5 años con los quintiles más altos y más bajos para el número de alelos de riesgo. Panel D: Riesgos a 5 años con deciles más altos y más bajos para el número de alelos de riesgo.

Figura 7. Riesgos de cáncer colorrectal en EE.UU. (mujeres) por categoría de edad, antecedentes familiares de cáncer colorrectal (pariente de primer grado) y por número de alelos de riesgo. Panel A: riesgos acumulados hasta los 70 años con los quintiles más altos y más bajos para el número de alelos de riesgo. Panel B: riesgos acumulados hasta los 70 años con deciles más altos y más bajos para el número de alelos de riesgo. Panel C: Riesgos a 5 años con los quintiles más altos y más bajos para el número de alelos de riesgo. Panel D: Riesgos a 5 años con deciles más altos y más bajos para el número de alelos de riesgo.

DESCRIPCIÓN DETALLADA DE LA INVENCIONTécnicas Generales y Definiciones Seleccionadas

- 5 **[0037]** A menos que se definan específicamente de otro modo, se entenderá que todos los términos técnicos y científicos utilizados en el presente documento tienen el mismo significado que los comúnmente entendidos por una persona con conocimientos ordinarios en la materia (p. ej., análisis del cáncer colorrectal, genética molecular, bioinformática y bioquímica).
- 10 **[0038]** A menos que se indique lo contrario, las técnicas moleculares y estadísticas utilizadas en la presente divulgación son procedimientos estándar, bien conocidos por los expertos en la materia. Dichas técnicas se describen y explican en toda la bibliografía en fuentes como, J. Perbal, A Practical Guide to Molecular Cloning, John Wiley and Sons (1984), J. Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbour Laboratory Press (1989), T.A. Brown (editor), Essential Molecular Biology: A Practical Approach, D.M. Glover y B.D. Hames (editores), DNA Cloning: A Practical Approach, volúmenes 1-4, IRL Press (1995 y 1996), y F.M. Ausubel et al. (editores), Current Protocols in Molecular Biology, Greene Pub. Associates y Wiley-Interscience (1988, incluidas todas las actualizaciones hasta la actualidad), Ed Harlow y David Lane (editores) Antibodies: A Laboratory Manual, Cold Spring Harbour Laboratory, (1988), y J.E. Coligan et al. (editores) Current Protocols in Immunology, John Wiley & Sons (incluidas todas las actualizaciones hasta la fecha).
- 20 **[0039]** Debe entenderse que esta invención no se limita a realizaciones particulares, que pueden, por supuesto, variar. También debe entenderse que la terminología empleada en el presente documento tiene por objeto describir únicamente realizaciones concretas y no pretende ser limitativa. Tal como se utilizan en esta especificación y en las reivindicaciones anexas, los términos en singular y las formas singulares "un/una", por ejemplo, incluyen opcionalmente referentes plurales a menos que el contenido indique claramente lo contrario. Así, por ejemplo, la referencia a "una sonda" incluye opcionalmente una pluralidad de moléculas de sonda; del mismo modo, dependiendo del contexto, el uso del término "un ácido nucleico" incluye opcionalmente, en la práctica, muchas copias de esa molécula de ácido nucleico.
- 25 **[0040]** Tal como se utiliza aquí, el término "aproximadamente", a menos que se indique lo contrario, se refiere a +/- 10%, más preferiblemente +/- 5%, más preferiblemente +/- 1%, del valor designado.
- 30 **[0041]** El término "y/o", p. ej., "X y/o Y" se entenderá en el sentido de "X e Y" o "X o Y" y se considerará que proporciona apoyo explícito para ambos significados o para cualquiera de ellos.
- 35 **[0042]** Tal como se utiliza en el presente documento, el término "cáncer colorrectal" abarca cualquier tipo de cáncer que pueda desarrollarse en el colon o el recto de un sujeto. Los términos "cáncer colorrectal", "cáncer de colon", "cáncer rectal" y "cáncer de intestino" pueden utilizarse indistintamente en el contexto de la presente divulgación.
- 40 **[0043]** Por ejemplo, el cáncer colorrectal puede caracterizarse como T estadio 1 - 4. En otro ejemplo, el cáncer colorrectal puede caracterizarse como estadio A - D de Dukes
- 45 **[0044]** Tal como se utiliza aquí, "cáncer colorrectal" también abarca un fenotipo que muestra una predisposición a desarrollar cáncer colorrectal en un individuo. Un fenotipo que muestre una predisposición al cáncer colorrectal puede, por ejemplo, mostrar una mayor probabilidad de que el cáncer se desarrolle en un individuo con el fenotipo que en los miembros de una población general relevante bajo un conjunto dado de condiciones ambientales (dieta, régimen de actividad física, localización geográfica, etc.). Por ejemplo, el cáncer colorrectal puede clasificarse clínicamente como premaligno (por ejemplo, hiperplasia, adenoma).
- 50 **[0045]** Un "polimorfismo" es un locus que es variable; es decir, dentro de una población, la secuencia de nucleótidos en un polimorfismo tiene más de una versión o alelo. Un ejemplo de polimorfismo es el "polimorfismo de nucleótido único", que es un polimorfismo en una única posición de nucleótido en un genoma (el nucleótido en la posición especificada varía entre individuos o poblaciones).
- 55 **[0046]** Tal como se utiliza aquí, el término "SNP" o "polimorfismo de nucleótido único" se refiere a una variación genética entre individuos; p. ej., una única posición de base nitrogenada en el ADN de los organismos que es variable. En el presente documento, "SNP" es el plural de SNP. Por supuesto, cuando uno se refiere al ADN en el presente documento, dicha referencia puede incluir derivados del ADN como amplicóns, transcritos de ARN del mismo, etc.
- 60 **[0047]** El término "alelo" se refiere a una de dos o más secuencias nucleotídicas diferentes que se producen o están codificadas en un locus específico, o a dos o más secuencias polipeptídicas diferentes codificadas por dicho locus. p. ej., un primer alelo puede ocurrir en un cromosoma, mientras que un segundo alelo ocurre en un segundo cromosoma homólogo, p. ej., como ocurre en diferentes cromosomas de un individuo heterocigoto, o entre diferentes individuos homocigotos o heterocigotos de una población. Un alelo se correlaciona "positivamente" con un rasgo cuando está ligado a él y cuando la presencia del alelo es un indicador de que el rasgo o la forma del rasgo se producirá en un individuo que comprenda el alelo. Un alelo se correlaciona "negativamente" con un rasgo cuando está ligado a él y cuando la presencia del alelo es un indicador de que un rasgo o forma de rasgo no se producirá en un individuo que comprenda el alelo. El
- 65

término "alelo de riesgo" se utiliza en el contexto de la presente divulgación para referirse a un alelo que indica una propensión genética a la susceptibilidad al cáncer colorrectal. Un sujeto puede ser homocigoto, heterocigoto o nulo para un alelo de riesgo concreto.

[0048] Un polimorfismo o alelo marcador está "correlacionado" o "asociado" con un fenotipo específico (susceptibilidad al cáncer colorrectal, etc.) cuando puede relacionarse estadísticamente (positiva o negativamente) con el fenotipo. Los métodos para determinar si un polimorfismo o alelo está vinculado estadísticamente son conocidos por los expertos en la materia. Es decir, el polimorfismo o polimorfismos especificados se dan con más frecuencia en una población de casos (p. ej., pacientes con cáncer colorrectal) que en una población de control (p. ej., individuos que no padecen cáncer colorrectal). A menudo se infiere que esta correlación es de naturaleza causal, pero no tiene por qué serlo: el simple vínculo genético con (asociación con) un locus para un rasgo que subyace al fenotipo es suficiente para que se produzca la correlación/asociación.

[0049] La expresión "desequilibrio de ligamiento" (LD) se utiliza para describir la correlación estadística entre dos genotipos polimórficos vecinos. Normalmente, la LD se refiere a la correlación entre los alelos de un gameto aleatorio en los dos loci, suponiendo un equilibrio de Hardy-Weinberg (independencia estadística) entre gametos. La LD se cuantifica con el parámetro de asociación de Lewontin (D') o con el coeficiente de correlación de Pearson (r) (Devlin y Risch, 1995). Dos loci con un valor LD de 1 se dice que están en LD completa. En el otro extremo, dos loci con un valor LD de 0 se denominan en equilibrio de ligamiento. El desequilibrio de ligamiento se calcula aplicando el algoritmo de maximización de expectativas (EM) para la estimación de frecuencias de haplotipos (Slatkin y Excoffier, 1996). Los valores de LD según la presente divulgación para genotipos/loci vecinos se seleccionan por encima de 0,5, más preferiblemente, por encima de 0,6, aún más preferiblemente, por encima de 0,7, preferiblemente, por encima de 0,8, más preferiblemente por encima de 0,9, idealmente aproximadamente de 1,0. Muchos de los SNP en desequilibrio de ligamiento con los SNP de la presente divulgación que se describen aquí tienen valores de LD de 0,9 o 1.

[0050] Otra forma en la que un experto en la materia puede identificar fácilmente SNP en desequilibrio de ligamiento con los SNP aquí descritos es determinando la puntuación LOD para dos loci. LOD son las siglas en inglés de "logaritmo de las probabilidades", una estimación estadística de si es probable que dos genes, o un gen y un gen de una enfermedad, estén situados cerca el uno del otro en un cromosoma y, por tanto, sea probable que se hereden. Por lo general, una puntuación LOD de entre 2 y 3 o superior significa que dos genes están situados cerca el uno del otro en el cromosoma. Así, los valores LOD para genotipos/loci vecinos pueden seleccionarse al menos por encima de 2, al menos por encima de 3, al menos por encima de 4, al menos por encima de 5, al menos por encima de 6, al menos por encima de 7, al menos por encima de 8, al menos por encima de 9, al menos por encima de 10, al menos por encima de 20, al menos por encima de 30, al menos por encima de 40, al menos por encima de 50.

[0051] Los SNP en desequilibrio de ligamiento con los SNP aquí divulgados pueden tener una distancia de recombinación genética especificada menor o igual a unos 20 centimorgan (cM) o menos. Por ejemplo, 15 cM o menos, 10 cM o menos, 9 cM o menos, 8 cM o menos, 7 cM o menos, 6 cM o menos, 5 cM o menos, 4 cM o menos, 3 cM o menos, 2 cM o menos, 1 cM o menos, 0,75 cM o menos, 0,5 cM o menos, 0,25 cM o menos, o 0,1 cM o menos. Por ejemplo, dos loci ligados dentro de un único segmento cromosómico pueden sufrir recombinación entre sí durante la meiosis con una frecuencia menor o igual a aproximadamente el 20%, aproximadamente el 19%, aproximadamente el 18%, aproximadamente el 17%, aproximadamente el 16%, aproximadamente el 15%, aproximadamente el 14%, aproximadamente el 13%, aproximadamente el 12%, aproximadamente el 11%, aproximadamente el 10%, aproximadamente el 9%, aproximadamente el 8%, aproximadamente el 7%, aproximadamente el 6%, aproximadamente el 5%, aproximadamente el 4%, aproximadamente el 3%, aproximadamente el 2%, aproximadamente el 1%, aproximadamente el 0,75%, aproximadamente el 0,5%, aproximadamente el 0,25% o aproximadamente el 0,1% o menos.

[0052] Los SNP en desequilibrio de ligamiento con los SNP aquí divulgados pueden estar dentro de un radio de al menos 100 kb (que se correlaciona en humanos con aproximadamente 0,1 cM, dependiendo de la tasa de recombinación local), al menos 50 kb, al menos 20 kb o menos entre sí.

[0053] Un enfoque ejemplar para la identificación de marcadores sustitutos para un SNP particular implica una estrategia simple que presume que los SNP que rodean al SNP objetivo están en desequilibrio de ligamiento y, por lo tanto, pueden proporcionar información sobre la susceptibilidad a la enfermedad. Por lo tanto, los marcadores potencialmente sustitutos pueden identificarse a partir de bases de datos de acceso público, como HAPMAP, mediante la búsqueda de SNP que cumplan determinados criterios que la comunidad científica haya considerado adecuados para la selección de candidatos a marcadores sustitutos.

[0054] "Frecuencia alélica" se refiere a la frecuencia (proporción o porcentaje) en la que un alelo está presente en un locus dentro de un individuo, dentro de una línea o dentro de una población de líneas. Por ejemplo, para un alelo "A", los individuos diploides de genotipo "AA", "Aa" o "aa" tienen frecuencias alélicas de 1,0, 0,5 o 0,0, respectivamente. Se puede estimar la frecuencia alélica dentro de una línea o población (p. ej., casos o controles) promediando las frecuencias alélicas de una muestra de individuos de esa línea o población. Del mismo modo, se puede calcular la frecuencia alélica dentro de una población de líneas promediando las frecuencias alélicas de las líneas que componen la población.

[0055] El término "frecuencia alélica" puede utilizarse para definir la frecuencia alélica menor (MAF). El MAF se refiere a

la frecuencia con la que se produce el alelo menos común en una población determinada.

[0056] Un individuo es "homocigoto" si sólo tiene un tipo de alelo en un locus determinado (p. ej., un individuo diploide tiene una copia del mismo alelo en un locus para cada uno de los dos cromosomas homólogos). Un individuo es "heterocigoto" si más de un tipo de alelo está presente en un locus determinado (p. ej., un individuo diploide con una copia de cada uno de dos alelos diferentes). El término "homogeneidad" indica que los miembros de un grupo tienen el mismo genotipo en uno o más loci específicos. En cambio, el término "heterogeneidad" se utiliza para indicar que los individuos del grupo difieren en genotipo en uno o más loci específicos.

[0057] Un "locus" es una posición o región cromosómica. Por ejemplo, un locus polimórfico es una posición o región donde se localiza un ácido nucleico polimórfico, un determinante de rasgo, un gen o un marcador. En otro ejemplo, un "locus genético" es una localización (región) cromosómica específica en el genoma de una especie donde se puede encontrar un gen específico.

[0058] Un "marcador", "marcador molecular" o "ácido nucleico marcador" se refiere a una secuencia de nucleótidos o producto codificado de la misma (p. ej., una proteína) utilizado como punto de referencia al identificar un locus o un locus vinculado. Un marcador puede derivarse de una secuencia de nucleótidos genómica o de secuencias de nucleótidos expresadas (p. ej., de un ARN, ARNm, un ADNc, etc.), o de un polipéptido codificado. El término también se refiere a secuencias de ácido nucleico complementarias o flanqueantes de las secuencias marcadoras, como los ácidos nucleicos utilizados como sondas o pares de cebadores capaces de amplificar la secuencia marcadora. Una "sonda marcadora" es una secuencia o molécula de ácido nucleico que puede utilizarse para identificar la presencia de un locus marcador, p. ej., una sonda de ácido nucleico complementaria a una secuencia de locus marcador. Los ácidos nucleicos son "complementarios" cuando se hibridan específicamente en solución, p. ej., según las reglas de emparejamiento de bases de Watson-Crick. Un "locus marcador" es un locus que puede utilizarse para rastrear la presencia de un segundo locus ligado, p. ej., un locus ligado o correlacionado que codifica o contribuye a la variación poblacional de un rasgo fenotípico. Por ejemplo, un locus marcador puede utilizarse para supervisar la segregación de alelos en un locus, como un locus de rasgo cuantitativo (QTL), que están genética o físicamente vinculados al locus marcador. Así, un "alelo marcador", alternativamente un "alelo de un locus marcador" es una de una pluralidad de secuencias de nucleótidos polimórficos encontradas en un locus marcador en una población que es polimórfica para el locus marcador.

[0059] Se divulgan aquí loci marcadores correlacionados con un fenotipo de interés, p. ej., cáncer colorrectal. Se espera que cada uno de los marcadores identificados se encuentre en estrecha proximidad física y genética (dando lugar a un enlace físico y/o genético) con un elemento genético, p. ej., un QTL que contribuya al fenotipo pertinente. Los marcadores correspondientes a polimorfismos genéticos entre miembros de una población pueden detectarse mediante métodos bien establecidos en la técnica. Entre ellos se incluyen, p. ej., los métodos de amplificación de secuencias específicas basados en la PCR, la detección de polimorfismos de longitud de fragmentos de restricción (RFLP), la detección de marcadores isoenzimáticos, la detección de hibridación específica de alelos (ASH), la detección de extensión de nucleótido único, la detección de secuencias variables amplificadas del genoma, la detección de replicación de secuencias autosostenida, la detección de repeticiones de secuencias simples (SSR), la detección de polimorfismos de nucleótido único (SNP) o la detección de polimorfismos de longitud de fragmentos amplificados (AFLP).

[0060] El término "amplificar" en el contexto de la amplificación de ácidos nucleicos es cualquier proceso mediante el cual se producen copias adicionales de un ácido nucleico seleccionado (o una forma transcrita del mismo). Los métodos de amplificación típicos incluyen varios métodos de replicación basados en polimerasas, incluida la reacción en cadena de la polimerasa (PCR), métodos mediados por ligasas como la reacción en cadena de la ligasa (LCR) y métodos de amplificación basados en la ARN polimerasa (p. ej., por transcripción).

[0061] Un "amplicón" es un ácido nucleico amplificado, p. ej., un ácido nucleico que se produce amplificando un ácido nucleico molde mediante cualquier método de amplificación disponible (p. ej., PCR, LCR, transcripción o similar).

[0062] Un ácido nucleico especificado es "derivado de" un ácido nucleico dado cuando se construye usando la secuencia del ácido nucleico dado, o cuando el ácido nucleico especificado se construye usando el ácido nucleico dado.

[0063] Un "gen" es una o más secuencia(s) de nucleótidos en un genoma que en conjunto codifican una o más moléculas expresadas, p. ej., un ARN, o polipéptido. El gen puede incluir secuencias codificantes que se transcriben en ARN que luego puede traducirse en una secuencia polipeptídica, y puede incluir secuencias estructurales o reguladoras asociadas que ayudan a la replicación o expresión del gen.

[0064] Un "genotipo" es la constitución genética de un individuo (o grupo de individuos) en uno o más loci genéticos. El genotipo se define por el alelo o alelos de uno o varios loci conocidos del individuo, normalmente, la compilación de alelos heredados de sus progenitores.

[0065] Un "haplotipo" es el genotipo de un individuo en una pluralidad de loci genéticos de una única cadena de ADN. Normalmente, los loci genéticos descritos por un haplotipo están física y genéticamente vinculados, es decir, en la misma cadena cromosómica.

[0066] Un "conjunto" de marcadores, sondas o cebadores se refiere a una colección o grupo de marcadores sondas, cebadores, o los datos derivados de los mismos, utilizados para un propósito común (p. ej., evaluar el riesgo de un individuo de desarrollar cáncer colorrectal). Con frecuencia, los datos correspondientes a los marcadores, sondas o cebadores, o derivados de su uso, se almacenan en un soporte electrónico. Si bien cada uno de los miembros de un conjunto posee utilidad con respecto al propósito especificado, los marcadores individuales seleccionados del conjunto, así como los subconjuntos que incluyen algunos, pero no todos los marcadores, también son eficaces para lograr el propósito especificado.

[0067] Los polimorfismos y genes, y las correspondientes sondas marcadoras, amplicóns o cebadores descritos anteriormente pueden incorporarse en cualquier sistema del presente documento, ya sea en forma de ácidos nucleicos físicos, o en forma de instrucciones del sistema que incluyan información de secuencia para los ácidos nucleicos. Por ejemplo, el sistema puede incluir cebadores o amplicóns correspondientes a (o que amplifican una porción de) un gen o polimorfismo descrito en el presente documento. Como en los métodos anteriores, el conjunto de sondas o cebadores marcadores detecta opcionalmente una pluralidad de polimorfismos en una pluralidad de dichos genes o loci genéticos. Así, por ejemplo, el conjunto de sondas o cebadores marcadores detecta al menos un polimorfismo en cada uno de estos genes, o cualquier otro polimorfismo, gen o locus definido en el presente documento. Dicha sonda o cebador puede incluir una secuencia de nucleótidos de dicho polimorfismo o gen, o un ácido nucleico complementario del mismo, o un producto transcrito del mismo (p. ej., una forma de ARNn o ARNm producida a partir de una secuencia genómica, p. ej., por transcripción o empalme).

[0068] Tal como se utilizan en el presente documento, las "curvas características operativas del receptor" se refieren a un trazado gráfico de la sensibilidad frente a la especificidad. (1 - especificidad) para un sistema clasificador binario a medida que varía su umbral de discriminación. El ROC también puede representarse de forma equivalente trazando la fracción de verdaderos positivos (TPR = tasa de verdaderos positivos) frente a la fracción de falsos positivos (FPR = tasa de falsos positivos). También se conoce como curva característica de funcionamiento relativo, porque es una comparación de dos características de funcionamiento (TPR y FPR) a medida que cambia el criterio. El análisis ROC proporciona herramientas para seleccionar modelos posiblemente óptimos y descartar los subóptimos independientemente de (y antes de especificar) el contexto de costes o la distribución de clases. Los métodos de uso en el contexto de la divulgación serán claros para los expertos en la materia.

[0069] Tal y como se utiliza aquí, el término "combinar la evaluación del riesgo genético con la evaluación del riesgo clínico para obtener el riesgo" se refiere a cualquier análisis matemático adecuado que se base en los resultados de las dos evaluaciones. Por ejemplo, los resultados de la evaluación del riesgo clínico y de la evaluación del riesgo genético pueden sumarse, más preferiblemente multiplicarse.

[0070] Tal como se utilizan aquí, los términos "cribado rutinario del cáncer colorrectal" y "cribado más frecuente" son términos relativos, y se basan en una comparación con el nivel de cribado recomendado a un sujeto que no tiene riesgo identificado de desarrollar cáncer colorrectal. Por ejemplo, el cribado sistemático puede incluir el cribado de materia fecal oculta, la colonoscopia o la sigmoidoscopia cada uno o dos años. A continuación, se analizan otros intervalos de tiempo para el cribado sistemático.

Evaluación del riesgo genético

[0071] El método de la presente invención se refiere a la evaluación del riesgo de un sujeto de desarrollar cáncer colorrectal mediante la realización de una evaluación del riesgo genético.

[0072] La evaluación del riesgo genético puede realizarse analizando el genotipo del sujeto en dos o más loci de polimorfismos de nucleótido único. Por ejemplo, pueden detectarse al menos 28 polimorfismos de un solo nucleótido. En otros ejemplos, se detectan al menos 29, al menos 30, al menos 31, al menos 32, al menos 33, al menos 34, al menos 35, al menos 36, al menos 37, al menos 38, al menos 39, al menos 40, al menos 41, al menos 42, al menos 43, al menos 44 polimorfismos de nucleótido único. La evaluación del riesgo genético del método de la invención se realiza detectando al menos 45 polimorfismos de nucleótido único.

[0073] Como apreciará el destinatario experto, cada SNP que aumenta el riesgo de desarrollar cáncer colorrectal tiene una razón de posibilidades de asociación con el cáncer colorrectal superior a 1,0. En algunas realizaciones, ninguno de los polimorfismos tiene una razón de posibilidades de asociación con el cáncer colorrectal superior a 3 o superior a 4.

[0074] Los ejemplos de SNP que se detectan como parte de la evaluación del riesgo genético incluyen, pero no se limitan a, SNP seleccionados del grupo que consiste en rs72647484, rs10911251, rs6687758, 6691170, rs11903757, rs812481, rs35360328, rs10936599, rs3987, rs35509282, rs647161, rs1321311, rs16892766, rs6983267, rs10505477, rs7014346, rs719725, rs10904849, rs10795668, rs704017, rs11190164, rs1035209, rs12241008, rs174537, rs4246215, rs174550, rs1535, rs3824999, rs3802842, rs3217810, rs3217901, rs10774214, rs11169552, rs7136702, rs3184504, rs59336, rs73208120, rs1957636, rs4444235, rs11632715, rs16969681, rs9929218, rs16941835, rs744166, rs4939827, rs10411210, rs1800469, rs2241714, rs2423279, rs4813802, rs961253, rs6066825, rs4925386 o rs5934683. Como también se divulga en el presente documento, los SNP que se detectan pueden ser SNP en desequilibrio de ligamiento con uno o más de los de la lista anterior. En el método de la invención, los SNP detectados son los de la Tabla 1. Como

también se describe en el presente documento, al realizar la evaluación del riesgo genético se detectan al menos 28 SNP de la Tabla 1 o un polimorfismo de nucleótido único en desequilibrio de ligamiento con uno o más de ellos. Alternativamente, se detectan al menos 29, al menos 30, al menos 31, al menos 32, al menos 33, al menos 34, al menos 35, al menos 36, al menos 37, al menos 38, al menos 39, al menos 40, al menos 41, al menos 42, al menos 43, al menos 44 polimorfismos de nucleótido único de la Tabla 1 o un polimorfismo de nucleótido único en desequilibrio de ligamiento con uno o más de ellos. En el método de la invención, se detectan al menos 45 polimorfismos de nucleótido único de la Tabla 1.

Tabla 1. SNP asociados al cáncer colorrectal. La tabla indica la nomenclatura del SNP, el gen o genes más cercanos o dentro de la probable diana reguladora del SNP, el genotipo del alelo de riesgo notificado, la frecuencia del alelo de riesgo notificada en los controles, la asociación notificada con el cáncer colorrectal por alelo de riesgo (razón de posibilidades), el riesgo relativo familiar (FRR) atribuible al SNP y la proporción del log FRR debida al SNP. *Gen más cercano o probable diana reguladora del SNP. Los SNP en desequilibrio de ligamiento se muestran entre corchetes [].

Locus	Gen*	SNP	Alelo de riesgo	OR por alelo de riesgo	Frecuencia de alelo de riesgo	FRR	Proporción de log FRR
1p36.2	WNT4; CDC42	rs72647484	T	1.21	0.91	1.003	0.37%
1q25.3	LAMC1	rs10911251	A	1.05	0.54	1.0008	0.07%
1q41	DUSP10; CICP13	rs6687758; [rs6691170]	G	1.09	0.2	1.0012	0.15%
2q32.3	NABP1; MYO18; SDPR	rs11903757	C	1.06	0.39	1.003	0.37%
3p14.1	LRIG1	rs812481	G	1.09	0.58	1.0018	0.22%
3p22.1	RPT1; CTNNA1	rs35360326	A	1.14	0.16	1.0023	0.29%
3q26.2	MYNN; TERC	rs10936599	C	1.06	0.75	1.0011	0.14%

(continuación)

4q26	NDST3	rs3987	C	1.36	0.44	1.0235	2.87%
4q32.2	FSTL5	rs35509282	A	1.53	0.09	1.0149	1.83%
5q31.1	PITX1; H2AFY	rs647161	A	1.11	0.67	1.0324	0.30%
6p21.31	CDKN1A	rs1321311	A	1.1	0.23	1.0016	0.20%
8q23.3	EIF3H	rs16892766	C	1.25	0.07	1.0032	0.40%
8q24.21	CCAT2; MYC	rs6963267 [rs10505477, rs7014346]	G	1.21	0.52	1.0091	1.12%
9q24	TPD52L3; UHRF2	rs719725	A	1.19	0.37	1.0011	0.13%
10p13	CUBN	rs10904849	G	1.14	0.66	1.0037	0.46%
10p14	GATA3	rs10795668	G	1.12	0.67	1.0028	0.35%
10q22.3	ZMIZ1; AS1	rs704017	G	1.06	0.67	1.0006	0.10%
10q24.2	SLC25A28; ENTPD7; COX15; CUTC; ABCC2	rs11190164 [rs1035209]	G	1.09	0.29	1.0015	0.19%
10q25	VTG1A	rs12241006	C	1.13	0.09	1.0012	0.15%
11q12.2	FADS1; FEN1	11qhap ^a ; [rs174537, rs4246215, rs174550, rs1535]	G	1.4	0.57	1.0281	3.41%
11q13.4	POLD3	rs3824999	G	1.08	0.5	1.0015	0.18%
11q23.1	COLCA2	rs3802842	C	1.11	0.29	1.0022	0.28%
12p13.32	CCND2	rs3217810	T	1.2	0.16	1.0045	0.55%
12p13.32	CCND2	rs3217901	G	1.1	0.41	1.0022	0.27%
12p13.32	CCND2	rs10774214	T	1.09	0.38	1.0018	0.22%
12q13.13	DIP2B; ATF1	rs11169552	C	1.09	0.72	1.0015	0.16%
12q13.13	LARP4; DIP2B	rs7136702	T	1.06	0.35	1.0008	0.10%
12q24.12	SH2B3	rs3184504	C	1.09	0.53	1.0019	0.23%
12q24.21	TBX3	rs59336	T	1.09	0.48	1.0019	0.23%
12q24.22	NOS1	rs73208120	G	1.16	0.11	1.0021	0.26%
14q22.2	BMP4	rs1957636	T	1.08	0.4	1.0014	0.18%
14q22.2	BMP4	rs4444235	C	1.11	0.46	1.0027	0.33%
15q13.3	SCG5; GREM1	rs11632715	A	1.12	0.47	1.0032	0.39%
15q13.3	SCG5; GREM1	rs16969081	T	1.18	0.09	1.0022	0.28%
16q22.1	COH1	rs9929218	G	1.1	0.71	1.0019	0.23%
16q24.1	FOXJ1	rs16941835	C	1.15	0.21	1.0032	0.40%
17q21	STAT3	rs744166	G	1.27	0.55	1.0142	1.74%
18q21.1	SMAD7	rs4939627	T	1.16	0.52	1.0069	0.84%

(continuación)

19q13.11	RHPN2	rs10411210	C	1.15	0.9	1.0018	0.22%
19q13.2	TMEM91; TGFBI	19qhap^; [rs1800469, rs2241714]	G	1.16	0.49	1.0055	0.68%
20p12.3	FERMT1; BMP2	rs2423279	C	1.14	0.3	1.0036	0.44%
20p12.3	FERMT1; BMP2	rs4813802	G	1.09	0.36	1.0017	0.21%
20p12.3	FERMT1; BMP2	rs961253	A	1.12	0.36	1.003	0.36%
20q13.1	PREX1	rs6066825	A	1.09	0.64	1.0017	0.21%
20q13.33	LAMA5	rs4925386	C	1.08	0.68	1.0013	0.16%

[0075] Los polimorfismos de nucleótido único en desequilibrio de ligamiento con uno o más de los polimorfismos de nucleótido único seleccionados de la Tabla 1 pueden tener valores LD de al menos 0,5, al menos 0,6, al menos 0,7, al menos 0,8. Alternativamente, los polimorfismos de nucleótido único en desequilibrio de ligamiento pueden tener valores de LD de al menos 0,9. En otra alternativa, los polimorfismos de nucleótido único en desequilibrio de ligamiento pueden tener valores de LD de al menos 1.

[0076] Algunos polimorfismos de nucleótido único son más informativos que otros para una evaluación de riesgo concreta. Por ejemplo, la evaluación del riesgo genético puede comprender la detección de rs3987, rs35509282 y rs744166, o un polimorfismo de nucleótido único en desequilibrio de ligamiento con uno o más de ellos.

[0077] La evaluación del riesgo genético del método de la invención comprende detectar rs72647484, rs10911251, rs6687758, rs11903757, rs812481, rs35360328, rs10936599, rs3987, rs35509282, rs647161, rs1321311, rs16892766, rs6983267, rs719725, rs10904849, rs10795668, rs704017, rs11190164, rs12241008, 11qhap (alguno o todos los rs174537, rs4246215, rs174550, y rs1535), rs3824999, rs3802842, rs3217810, rs3217901, rs10774214, rs11169552, rs7136702, rs3184504, rs59336, rs73208120, rs1957636, rs4444235, rs11632715, rs16969681, rs9929218, rs16941835, rs744166, rs4939827, rs10411210, 19qhap^ (alguno o todos los rs1800469 y rs2241714), rs2423279, rs4813802, rs961253, rs6066825, rs4925386. Como también se describe en el presente documento, la evaluación del riesgo genético puede comprender la detección de un polimorfismo de nucleótido único en desequilibrio de ligamiento con uno o más de los SNP enumerados anteriormente.

[0078] La evaluación del riesgo genético del método de la invención comprende detectar la presencia del polimorfismo de nucleótido único rs5934683. Como también se describe en el presente documento, la evaluación del riesgo genético puede comprender la detección de un polimorfismo de nucleótido único en desequilibrio de ligamiento.

[0079] En algunas realizaciones, el número de SNP evaluados se basa en la mejora de reclasificación neta en la predicción de riesgo calculada utilizando el índice de reclasificación neta (NRI) (Pencina et al., 2008). En algunas realizaciones, la mejora de reclasificación neta de los métodos de la presente divulgación es superior a 0,01.

[0080] En otras realizaciones, la mejora de reclasificación neta de los métodos de la presente divulgación es superior a 0,05. En algunas realizaciones, la mejora de reclasificación neta de los métodos de la presente divulgación es superior a 0,1.

[0081] Los SNP en desequilibrio de ligamiento con los mencionados específicamente en el presente documento son fácilmente identificables por los expertos en la materia. Ejemplos de tales SNP incluyen cuatro SNP perfectamente correlacionados dentro de 11q12.2 (rs174537, rs4246215, rs174550, y rs1535). Estos cuatro SNP se denominan en la presente divulgación haplotipo 11q12.2. Otro ejemplo son los genes rs1800469 y rs2241714, que se encuentran en 19q13.2. Estos SNP también están perfectamente correlacionados y se denominan en la presente divulgación haplotipo 19q13.2. Otros ejemplos incluyen rs6687758 y rs6691170, localizados en 1q41; rs10505477, rs6983267 y rs7014346, localizados en 8q24.21; rs11632715 y rs16969681 localizados en 15q31; rs1035209, rs11190164 localizados en 10q24.2; rs11169552, rs7136702 localizados en 12q13.13 (más ejemplos posibles en la Tabla 2).

Tabla 2. Lista de SNP (SNP correlacionados) en LD* con los seis principales SNP de riesgo (DbSNP). Se muestran los SNP con un r^2 superior a 0,08 (poblaciones afroamericana, americana, asiática y europea) en el conjunto de datos HAPMAP (<http://hapmap.ncbi.nlm.nih.gov>).

DbSNP	Posición de DbSNP	SNP correlacionado	Posición de SNP correlacionado	r^2	D'
rs16892766	chr8:117630683	rs16888589	chr8:117635602	1	1
		rs11986063	chr8:117640315	0.85	0.98
rs35509282	chr4:163333405	rs11736440	chr4:163336693	0.99	1
		rs12508784	chr4:163333299	0.86	1
		rs12511058	chr4:163326723	0.84	1
		rs17042479	chr4:163325411	0.85	1
		rs17600575	chr4:163328335	0.85	1
		rs2122494	chr4:163331379	0.98	1
		rs57336275	chr4:163341215	0.98	1
		rs74964851	chr4:163338255	0.98	1
		rs79783178	chr4:163325957	0.88	1
		rs9998942	chr4:163340404	0.98	1
		rs12642547	chr4:163337313	0.85	0.99
		rs12645341	chr4:163337355	0.85	0.99
		rs59363334	chr4:163340796	0.85	0.99
		rs11100440	chr4:163324864	0.81	0.97
rs3967	chr4:118759055	rs10018600	chr4:118776858	0.99	1
		rs10026807	chr4:118761523	0.97	1
		rs10026879	chr4:118761446	0.87	1
		rs12643469	chr4:118775565	1	1
		rs4317266	chr4:118778909	0.99	1
		rs4597906	chr4:118758795	0.98	1
		rs5861370	chr4:118764485	0.94	1
		rs7676593	chr4:118763497	0.98	1

(continuación)

		rs7684690	chr4:118774949	0.93	1
		rs1459530	chr4:118746231	0.83	0.99
		rs1459528	chr4:118750348	0.85	0.99
		rs1459529	chr4:118750315	0.85	0.99
		rs1459531	chr4:118742872	0.82	0.99
		rs4240312	chr4:118734518	0.81	0.99
		rs4270637	chr4:118744735	0.82	0.99
		rs4382104	chr4:118752001	0.85	0.99
		rs4834639	chr4:118755142	0.82	0.99
		rs6852960	chr4:118741585	0.82	0.99
		rs4377656	chr4:118782785	0.81	0.98
		rs7685408	chr4:118752469	0.87	0.97
		rs12503813	chr4:118784946	0.86	0.96
		rs13147985	chr4:118788434	0.88	0.96
		rs151286737	chr4:118790567	0.87	0.96
		rs4353970	chr4:118752091	0.86	0.95
		rs6824201	chr4:118736905	0.83	0.93
		rs11098407	chr4:118733381	0.82	0.92
		rs11562851	chr4:118735934	0.82	0.92
		rs11562871	chr4:118733490	0.82	0.92
		rs1380373	chr4:118736995	0.82	0.92
		rs17885121	chr4:118733657	0.82	0.92
		rs11427328	chr4:118737132	0.82	0.92
		rs6856317	chr4:118784120	0.82	0.92
		rs4594794	chr4:118788352	0.82	0.91
		rs6823808	chr4:118787965	0.82	0.91
		rs70941133	chr4:118784105	0.81	0.91
rs6983267	chr8:128413305	rs10505474	chr8:128417504	0.84	1
		rs10808556	chr8:128413147	0.84	1
		rs10956366	chr8:128423491	0.83	1
		rs10956370	chr8:128424728	0.83	1
		rs11778075	chr8:128421128	0.84	1
		rs11784983	chr8:128421348	0.84	1
		rs11998706	chr8:128422098	0.84	1
		rs12678562	chr8:128422488	0.84	1
		rs2060778	chr8:128420117	0.84	1
		rs3847137	chr8:128414498	0.84	1
		rs3933712	chr8:128420265	0.84	1
		rs4278948	chr8:128427372	0.84	1
		rs4871022	chr8:128427720	0.84	1
		rs4871788	chr8:128421785	0.84	1
		rs4871789	chr8:128428061	0.84	1
		rs7013328	chr8:128423911	0.83	1
		rs7018387	chr8:128424883	0.82	1
		rs7018368	chr8:128424933	0.83	1
		rs7018371	chr8:128424899	0.82	1
		rs7837328	chr8:128423127	0.83	1
		rs7837626	chr8:128423341	0.83	1
		rs7837644	chr8:128423398	0.83	1
		rs7837706	chr8:128423184	0.83	1
		rs871135	chr8:128426393	0.84	1
		rs12682374	chr8:128410948	0.97	0.99
rs72647484	chr1:22587728	rs2744997	chr1:22583855	0.86	1

(continuación)

		rs2744742	chr1:22566927	0.83	1
		rs2744748	chr1:22573163	0.83	1
		rs2744752	chr1:22575306	0.83	1
		rs2744753	chr1:22576327	0.86	1
		rs2744754	chr1:22576467	0.86	1
		rs2744756	chr1:22578619	0.86	1
		rs2807329	chr1:22565060	0.83	1
		rs2807332	chr1:22566847	0.96	1
		rs2807334	chr1:22566696	0.96	1
		rs2807335	chr1:22573764	0.96	1
		rs2807340	chr1:22580473	0.81	1
		rs28617726	chr1:22585280	1	1
		rs72647481	chr1:22584718	0.86	1
		rs72647481	chr1:22584718	1	1
		rs72647483	chr1:22587009	0.86	1
		rs72647483	chr1:22587009	1	1
		rs72647488	chr1:22590009	0.81	1
		rs72647488	chr1:22590009	0.89	1
		rs72647489	chr1:22590125	0.81	1
		rs72647489	chr1:22590125	0.89	1
		rs2744723	chr1:22535268	0.85	0.92
rs744166	chr17:40514201	rs1026916	chr17:40529835	0.88	1
		rs11079043	chr17:40545770	0.93	1
		rs11440924	chr17:40517857	0.99	1
		rs12601611	chr17:40497828	0.93	1
		rs12602466	chr17:40511946	0.9	1
		rs12937642	chr17:40525760	0.92	1
		rs12942547	chr17:40527544	0.85	1
		rs12942611	chr17:40535184	1	1
		rs12943176	chr17:40496447	0.93	1
		rs12949818	chr17:40526273	0.81	1
		rs12950549	chr17:40496594	1	1
		rs13342031	chr17:40536871	0.93	1
		rs17864075	chr17:40541608	1	1
		rs17884090	chr17:40518396	1	1
		rs17885629	chr17:40525098	0.81	1
		rs17885741	chr17:40498944	1	1
		rs17886724	chr17:40496163	1	1
		rs1905340	chr17:40520390	0.93	1
		rs1905341	chr17:40520597	0.9	1
		rs2306581	chr17:40500265	1	1
		rs35314169	chr17:40515626	0.93	1
		rs35840966	chr17:40521204	1	1
		rs35901220	chr17:40528168	0.94	1
		rs35950688	chr17:40499198	1	1
		rs3736161	chr17:40497835	1	1
		rs3736162	chr17:40497839	0.92	1
		rs3736164	chr17:40539825	0.93	1
		rs3785896	chr17:40515120	0.93	1
		rs3816769	chr17:40498273	0.99	1
		rs3869540	chr17:40492540	0.9	1
		rs4103200	chr17:40507065	0.93	1
		rs4795647	chr17:40543992	0.91	1

(continuación)

5	rs4796791	chr17:40530763	1	1
	rs58288833	chr17:40496701	0.9	1
	rs61454571	chr17:40538298	0.89	1
	rs62075772	chr17:40504250	1	1
	rs6503695	chr17:40499533	0.93	1
	rs6503696	chr17:40499804	0.93	1
10	rs6503697	chr17:40501579	0.93	1
	rs7211777	chr17:40534075	1	1
	rs7214610	chr17:40521787	0.92	1
	rs7216516	chr17:40517875	0.83	1
15	rs7217555	chr17:40496024	1	1
	rs7219059	chr17:40521670	0.92	1
	rs7219739	chr17:40531761	1	1
	rs7224007	chr17:40528786	0.92	1
20	rs7224416	chr17:40528702	0.92	1
	rs8068748	chr17:40532701	1	1
	rs8069545	chr17:40494902	0.92	1
	rs8070763	chr17:40536396	1	1
25	rs8071537	chr17:40530895	1	1
	rs8072391	chr17:40495380	1	1
	rs8073517	chr17:40503324	1	1
	rs8073636	chr17:40525719	0.99	1
	rs8075676	chr17:40505202	0.93	1
30	rs8076051	chr17:40505134	1	1
	rs8081037	chr17:40499158	0.91	1
	rs957970	chr17:40519860	1	1
	rs957971	chr17:40519925	1	1
35	rs9891119	chr17:40507980	1	1
	rs9895473	chr17:40515722	0.93	1
	rs9897369	chr17:40523725	0.86	1
	rs9912773	chr17:40510534	0.92	1
40	rs9913597	chr17:40510316	1	1
	rs35455295	chr17:40496438	0.95	1
	rs3869550	chr17:40492687	0.96	1
	rs4796793	chr17:40542210	0.92	0.99
45	rs11328125	chr17:40537526	0.91	0.98
	rs10706259	chr17:40492373	0.83	0.97
	rs2354155	chr17:40546652	0.84	0.95
	rs35561964	chr17:40536575	0.82	0.96
	rs34972443	chr17:40502074	0.83	0.93
50	rs2128786	chr17:40547327	0.81	0.91

Evaluación Clínica de Riesgos

55 **[0082]** Los métodos de la invención y tal como se divulgan aquí pueden comprender la realización de una evaluación del riesgo clínico del sujeto. Los resultados de la evaluación del riesgo clínico pueden combinarse con la evaluación del riesgo genético para obtener el riesgo del sujeto de desarrollar cáncer colorrectal.

60 **[0083]** En el método de la invención puede utilizarse cualquier procedimiento adecuado de evaluación del riesgo clínico. Preferiblemente, la evaluación del riesgo clínico no implica el genotipado del sujeto en uno o más loci. No obstante, el procedimiento de evaluación del riesgo clínico puede incluir la obtención de información sobre mutaciones en los genes *MLH1*, *MSH2* y *MSH6* y el estado de inestabilidad de microsatélites.

65 **[0084]** En algunas realizaciones, el procedimiento de evaluación del riesgo clínico incluye la obtención de información del sujeto sobre uno o más de los siguientes aspectos: antecedentes médicos de cáncer colorrectal y/o pólipos, edad, antecedentes familiares de cáncer colorrectal y/o pólipos y/u otro tipo de cáncer, incluida la edad del familiar en el momento

- del diagnóstico, resultados de colonoscopias y/o sigmoidoscopias previas, resultados de pruebas previas de sangre oculta en heces, peso, índice de masa corporal, altura, sexo, antecedentes de consumo de alcohol, antecedentes de tabaquismo, antecedentes de ejercicio, dieta (p. ej., consumo de folato, verduras, carne roja, frutas, fibra y grasas saturadas), prevalencia de enfermedad inflamatoria intestinal, raza/etnia, uso de aspirina y NSAID, implantación de la sustitución estrogénica y uso de anticonceptivos orales. Por ejemplo, el procedimiento de evaluación del riesgo clínico puede incluir la obtención de información del sujeto sobre los antecedentes de cáncer colorrectal de un familiar de primer grado. En otro ejemplo, el procedimiento de evaluación del riesgo clínico incluye la obtención de información del sujeto sobre la edad y/o los antecedentes de cáncer colorrectal de un familiar de primer grado.
- [0085]** En algunas realizaciones, la evaluación del riesgo clínico incluye detalles relativos a los antecedentes familiares de cáncer colorrectal de al menos algunos, preferiblemente todos, los familiares de primer grado.
- [0086]** En algunas realizaciones, la historia familiar de cáncer colorrectal implica un análisis de la historia familiar multigeneracional. Tal y como se utiliza aquí, "historia familiar multigeneracional" se refiere al análisis de 2 o más generaciones. La historia familiar multigeneracional puede incluir un análisis de, por ejemplo, entre la misma generación (por ejemplo, primos), y/o entre generaciones (por ejemplo, tíos y tías). Por ejemplo, en algunas realizaciones, la evaluación del riesgo clínico incluye detalles relativos a los antecedentes familiares de cáncer colorrectal de al menos algunos, preferiblemente todos, los familiares de segundo grado. En algunas realizaciones, la evaluación del riesgo clínico incluye detalles relativos a los antecedentes familiares de cáncer colorrectal de al menos algunos, preferiblemente todos, los familiares de segundo y tercer grado.
- [0087]** En algunas realizaciones, el procedimiento de evaluación del riesgo clínico proporciona una estimación del riesgo de que el sujeto desarrolle cáncer colorrectal durante el siguiente periodo de 5 años (es decir, riesgo a 5 años). En un ejemplo, el riesgo a 5 años determinado por la evaluación del riesgo clínico se sitúa entre el 1% y el 3% aproximadamente. En otro ejemplo, el riesgo a 5 años determinado por la evaluación del riesgo clínico se sitúa entre el 1,5% y el 2% aproximadamente.
- [0088]** En algunas realizaciones, el procedimiento de evaluación del riesgo clínico proporciona una estimación del riesgo de que el sujeto desarrolle cáncer colorrectal durante el próximo periodo de 10 años (es decir, riesgo a 10 años). En un ejemplo, el riesgo a 10 años determinado por la evaluación del riesgo clínico se sitúa entre el 1% y el 3% aproximadamente. En otro ejemplo, el riesgo a 5 años determinado por la evaluación del riesgo clínico se sitúa entre el 1,5% y el 2% aproximadamente.
- [0089]** En algunas realizaciones, el procedimiento de evaluación del riesgo clínico proporciona una estimación del riesgo de que el sujeto desarrolle cáncer colorrectal hasta los 70 años (es decir, el riesgo de por vida). En un ejemplo, el riesgo a lo largo de la vida determinado por la evaluación del riesgo clínico se sitúa entre el 15% y el 30% aproximadamente. En otro ejemplo, el tiempo de vida determinado por la evaluación del riesgo clínico se sitúa entre el 20% y el 25% aproximadamente.
- [0090]** En algunas realizaciones, la evaluación del riesgo clínico utiliza un modelo que calcula el riesgo absoluto de desarrollar cáncer de colon. Por ejemplo, el riesgo absoluto de desarrollar cáncer de colon puede calcularse utilizando las tasas de incidencia del cáncer y teniendo en cuenta el riesgo concurrente de morir por otras causas distintas del cáncer de colon. En algunas realizaciones, la evaluación del riesgo clínico proporciona un riesgo absoluto a 5 años de desarrollar cáncer de colon. En algunas realizaciones, la evaluación del riesgo clínico proporciona un riesgo absoluto a 10 años de desarrollar cáncer de colon.
- [0091]** Ejemplos de procedimientos de evaluación del riesgo clínico incluyen, entre otros, el Índice de Riesgo de Cáncer de Harvard, la Herramienta de Evaluación del Riesgo de Cáncer Colorrectal del Instituto Nacional del Cáncer, la Herramienta de la Clínica Cleveland, el modelo de probabilidad de Reparación de Desajustes (también conocido como MMRpro), la Herramienta de Predicción del Riesgo Colorrectal (CRiPT) y similares (véase, por ejemplo, Usher-Smith et al., 2015). En estos algoritmos ejemplares de predicción del riesgo se ha recopilado un amplio conjunto de investigaciones centradas en las mutaciones de alto riesgo y en los factores de riesgo fenotípicos.
- [0092]** El Índice de Riesgo de Cáncer de Harvard predice el riesgo a 10 años de desarrollar cáncer de colon utilizando datos de antecedentes familiares (parientes de primer grado con cáncer de colon) y factores ambientales como el índice de masa corporal, el uso de aspirina, el consumo de cigarrillos, los antecedentes de enfermedad inflamatoria intestinal, la altura, la actividad física, la sustitución de estrógenos, el uso de anticonceptivos orales y el consumo de folato, verduras, alcohol, carne roja, frutas, fibra y grasas saturadas. En un ejemplo, el procedimiento de evaluación del riesgo clínico utiliza el Índice de Riesgo de Cáncer de Harvard para predecir el riesgo a 10 años de que el sujeto desarrolle cáncer de colon.
- [0093]** La Herramienta de Evaluación del Riesgo de Cáncer Colorrectal predice los riesgos a 5, 10, 20 años y de por vida de desarrollar cáncer colorrectal en personas mayores de 50 años en función de la edad, el sexo, el uso de sigmoidoscopia y/o colonoscopia, la actividad actual en el tiempo libre, el uso de aspirina y NSAID, los antecedentes de tabaquismo, el índice de masa corporal, los antecedentes de reemplazo hormonal y el consumo de verduras. En un ejemplo, el procedimiento de evaluación del riesgo clínico utiliza la Herramienta de Evaluación del Riesgo de Cáncer Colorrectal para predecir el riesgo a 5 años de que el sujeto desarrolle cáncer colorrectal. En otro ejemplo, el procedimiento de evaluación

del riesgo clínico utiliza la Herramienta de Evaluación del Riesgo de Cáncer Colorrectal para predecir el riesgo a 10 años de que el sujeto desarrolle cáncer colorrectal. En otro ejemplo, el procedimiento de evaluación del riesgo clínico utiliza la Herramienta de Evaluación del Riesgo de Cáncer Colorrectal para predecir el riesgo a 20 años de que el sujeto desarrolle cáncer colorrectal. En otro ejemplo, el procedimiento de evaluación del riesgo clínico utiliza la Herramienta de Evaluación del Riesgo de Cáncer Colorrectal para predecir el riesgo a lo largo de la vida de que el sujeto desarrolle cáncer colorrectal.

[0094] La herramienta de la Clínica Cleveland proporciona una puntuación de riesgo de cáncer colorrectal basada en la edad, el sexo, el origen étnico, el peso, la altura, el uso de sigmoidoscopia y/o colonoscopia, la prueba de sangre oculta en heces, el tabaquismo, el ejercicio, los antecedentes de cáncer colorrectal y pólipos, y el consumo de verduras y frutas.

[0095] El modelo MMRpro predice los riesgos a cinco años y de por vida de desarrollar cáncer colorrectal y de endometrio basándose en mutaciones en los genes *MLH1*, *MSH2* y *MSH6*, así como en factores ambientales como los antecedentes familiares de la enfermedad, el estado de inestabilidad de microsátélites, la edad y el origen étnico. En un ejemplo, el procedimiento de evaluación del riesgo clínico utiliza el modelo MMRpro para predecir el riesgo a 5 años de que el sujeto desarrolle cáncer colorrectal. En otro ejemplo, el procedimiento de evaluación del riesgo clínico utiliza el modelo MMRpro para predecir el riesgo a lo largo de la vida de que el sujeto desarrolle cáncer colorrectal.

[0096] El modelo de la Herramienta de Predicción del Riesgo Colorrectal (CRiPT) utiliza la historia familiar multigeneracional mediante un modelo poligénico mixto de genes principales para estimar el riesgo de cáncer colorrectal.

Cálculo del Riesgo Relativo del SNP Compuesto "Riesgo Genético"

[0097] El "riesgo genético" de un individuo puede definirse como el producto de los valores de riesgo relativo del genotipo para cada SNP evaluado. Se puede utilizar un modelo de riesgo log-aditivo para definir tres genotipos AA, AB y BB para un único SNP con valores de riesgo relativo de 1, OR y OR², bajo un modelo de enfermedad rara, donde OR es la razón de posibilidades de enfermedad previamente descrita para el alelo de alto riesgo, B, frente al alelo de bajo riesgo, A. Si el alelo B tiene frecuencia (p), entonces estos genotipos tienen frecuencias poblacionales de (1 - p)², 2p(1 - p), y p², asumiendo el equilibrio de Hardy-Weinberg. Los valores de riesgo relativo del genotipo para cada SNP pueden entonces escalarse de forma que, basándose en estas frecuencias, el riesgo relativo medio en la población sea 1. En concreto, dado el riesgo relativo medio poblacional sin escalar:

$$(\mu) = (1 - p)^2 + 2p(1 - p)OR + p^2OR^2$$

[0098] Los valores de riesgo ajustados 1/μ, OR/μ y OR²/μ se utilizan para los genotipos AA, AB y BB. A los genotipos ausentes se les asigna un riesgo relativo de 1. La siguiente fórmula puede utilizarse para definir el riesgo genético:

SNP₁ x SNP₂ x SNP₃ x SNP₄ x SNP₅ x SNP₆ x SNP₇, x SNP, etc.
Se pueden realizar cálculos similares para polimorfismos no SNP.

[0099] Un método alternativo para calcular el riesgo SNP compuesto se describe en Mavaddat et al. (2015). En este ejemplo, se utiliza la siguiente fórmula;

$$PRS = \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \beta_n x_n$$

donde β_k es el log razón de posibilidades (OR) por alelo para el cáncer de colon asociado con el alelo menor para el SNP k, y x_k el número de alelos para el mismo SNP (0, 1 o 2), n es el número total de SNP y PRS es la puntuación de riesgo poligénico (que también puede denominarse riesgo SNP compuesto).

[0100] Se prevé que el "riesgo" de un sujeto humano de desarrollar cáncer colorrectal pueda proporcionarse como un riesgo relativo (o cociente de riesgos) o un riesgo absoluto, según sea necesario.

[0101] En algunas realizaciones, la evaluación del riesgo genético obtiene el "riesgo relativo" de un sujeto humano de desarrollar cáncer colorrectal. El riesgo relativo (o cociente de riesgos), medido como la incidencia de una enfermedad en individuos con una característica (o exposición) concreta dividida por la incidencia de la enfermedad en individuos sin la característica, indica si esa exposición concreta aumenta o disminuye el riesgo. El riesgo relativo es útil para identificar las características que se asocian a una enfermedad, pero por sí mismo no es especialmente útil para orientar las decisiones de cribado, ya que la frecuencia del riesgo (incidencia) se anula.

[0102] En algunas realizaciones, la evaluación del riesgo genético obtiene el "riesgo absoluto" de un sujeto humano de desarrollar cáncer colorrectal. El riesgo absoluto es la probabilidad numérica de que un sujeto humano desarrolle cáncer colorrectal en un periodo determinado (por ejemplo, 5, 10, 15, 20 o más años). Refleja el riesgo de un sujeto humano de desarrollar cáncer colorrectal en la medida en que no considera diversos factores de riesgo de forma aislada.

Evaluación Clínica Combinada x Riesgo Genético

[0103] Al combinar la evaluación del riesgo clínico con la evaluación del riesgo genético para obtener el "riesgo" de un sujeto humano de desarrollar cáncer colorrectal, puede utilizarse la siguiente fórmula: [Riesgo (es decir, Evaluación Clínica x Riesgo SNP)] = [Riesgo de Evaluación Clínica] x SNP₁ x SNP₂ x SNP₃ x SNP₄ x SNP₅ x SNP₆ x SNP₇ x SNP₈,... x SNP₄₅ etc.

[0104] Donde Evaluación Clínica es el riesgo proporcionado por la evaluación clínica, y SNP₁ a SNP₄₅ son el riesgo relativo para los SNP individuales, cada uno escalado para tener una media poblacional de 1 como se ha descrito anteriormente. Dado que los valores de riesgo SNP se han "centrado" para tener un riesgo medio poblacional de 1, si se asume la independencia entre los SNP, entonces el riesgo medio poblacional a través de todos los genotipos para el valor combinado es coherente con la estimación de riesgo subyacente de la Evaluación Clínica.

[0105] En algunas realizaciones, el riesgo de un sujeto humano de desarrollar cáncer colorrectal se calcula mediante [Riesgo de evaluación clínica] x SNP₁ x SNP₂ x SNP₃ x SNP₄ x SNP₅ x SNP₆ x SNP₇ x SNP₈,... x SNP₄₅ etc. En algunas realizaciones, el riesgo de un sujeto humano de desarrollar cáncer colorrectal se calcula mediante [Evaluación clínica del riesgo a 5 años] x SNP₁ x SNP₂ x SNP₃ x SNP₄ x SNP₅ x SNP₆ x SNP₇ x SNP₈,... x SNP₄₅ etc.

[0106] En algunas realizaciones, el riesgo de un sujeto humano de desarrollar cáncer colorrectal se calcula mediante [Evaluación clínica riesgo de por vida] x SNP₁ x SNP₂ x SNP₃ x SNP₄ x SNP₅ x SNP₆ x SNP₇ x SNP₈,... x SNP₄₅ etc. En algunas realizaciones, la Evaluación Clínica se lleva a cabo evaluando uno o más de los siguientes aspectos: historial médico de cáncer colorrectal, edad, historial familiar de cáncer colorrectal, resultados de colonoscopias/sigmoidoscopias previas y raza/etnia para proporcionar un riesgo clínico. En esta realización, el riesgo (es decir, el riesgo genético combinado x el riesgo clínico) viene dado por:

[Riesgo (es decir, riesgo clínico x genético)] = $\text{factor clínico}_1 \times \text{factor clínico}_2 \dots \times \text{factor clínico}_5$
factor] x SNP₁ x SNP₂ x SNP₃ x SNP₄ x SNP₅ x SNP₆ x SNP₇ x SNP₈,... x SNP₄₅ etc.

[0107] En algunas realizaciones, la Evaluación Clínica se lleva a cabo evaluando el historial de familiares de primer grado de cáncer colorrectal para proporcionar un riesgo clínico. En esta realización, el riesgo (es decir, el riesgo genético combinado x el riesgo clínico) viene dado por:

[Riesgo (es decir, riesgo clínico x genético)] = [Riesgo clínico asociado con tener un pariente de primer grado con cáncer colorrectal] x SNP₁ x SNP₂ x SNP₃ x SNP₄ x SNP₅ x SNP₆ x SNP₇ x SNP₈,... x SNP₄₅ etc.

[0108] En algunas realizaciones, se puede estimar la proporción del riesgo relativo familiar logarítmico (FRR; la razón de posibilidades para el cáncer colorrectal asociada con tener un pariente de primer grado con cáncer colorrectal) que podría ser atribuible a los alelos de riesgo de los SNP (suponiendo la detección de 45 SNP, equilibrio de Hardy-Weinberg para cada SNP, equilibrio de vinculación entre los SNP y un modelo multiplicativo para las asociaciones de los SNP con el riesgo de cáncer colorrectal). SNP₁,... SNP₄₅ son SNP de la Tabla 1 y clínico₄₆,... clínico, son factores clínicos (nota: podría tratarse de cualquier factor hereditario que contribuya al FRR). Entonces, si G_i es una variable aleatoria que da el número de alelos de riesgo en SNP_i para una persona aleatoria de la población, entonces G₁,..., G_m son todas variables aleatorias independientes (por equilibrio de ligamiento) y la log-razón de probabilidades para una persona aleatoria es X₁ + ... + X_m (por modelo multiplicativo asumido), donde X_i = G_i log OR_i and OR_i es la razón de probabilidades por alelo para SNP_i. Una fórmula de Antoniou et al. 2003 ha derivado rigurosamente en Win et al. 2014 se convierten en log FRR = $\frac{1}{2}[\text{Var}(X_1) + \dots + \text{Var}(X_m)]$. Esto demuestra que el log FRR es la suma de componentes independientes de los SNP conocidos y desconocidos asociados al cáncer colorrectal. La proporción del log FRR debida a los SNP conocidos es $\frac{1}{2}[\text{Var}(X_1) + \dots + \text{Var}(X_{45})]/\text{log FRR}$, mientras que la proporción debida a factor(es) clínico(s) es uno menos este valor. En caso necesario, pueden incorporarse otros factores clínicos al cálculo anterior.

[0109] En algunas realizaciones, la evaluación del riesgo genético se combina con la evaluación del riesgo clínico para obtener el "riesgo relativo" de un sujeto humano de desarrollar cáncer colorrectal. En algunas realizaciones, la evaluación del riesgo genético se combina con la evaluación del riesgo clínico para obtener el "riesgo absoluto" de un sujeto humano de desarrollar cáncer colorrectal.

Sujetos

[0110] El término "sujeto", tal como se utiliza aquí, se refiere a un sujeto humano. Términos como "sujeto", "paciente" o "individuo" son términos que, en su contexto, pueden utilizarse indistintamente en la presente divulgación. En un ejemplo, los métodos de la invención pueden utilizarse para el cribado rutinario de sujetos. El cribado rutinario puede incluir la realización de pruebas a los sujetos a intervalos de tiempo predeterminados. Los intervalos de tiempo ejemplares incluyen el cribado mensual, trimestral, semestral, anual, cada dos años o cada tres años.

- 5 **[0111]** Los datos de riesgo actuales sugieren que la persona media alcanza el umbral de riesgo para el cribado de sangre oculta en heces (que recomiendan la mayoría de los programas nacionales de cribado) en torno a los 50 años de edad. Sin embargo, los presentes inventores han descubierto utilizando los métodos de la presente divulgación que algunos individuos deberían someterse a un cribado de sangre oculta en heces mucho antes de alcanzar los 50 años de edad, en particular si a un pariente de primer grado de estos sujetos se le ha diagnosticado cáncer colorrectal. Estos resultados sugieren que los sujetos menores de 50 años deben ser evaluados utilizando los métodos de la presente divulgación. Por consiguiente, en un ejemplo, los sujetos examinados mediante los métodos de la presente divulgación tienen al menos 38, al menos 39, al menos 40, al menos 41, al menos 42, al menos 43, al menos 44, al menos 45, al menos 46, al menos 47, al menos 48, al menos 49 años de edad. En un ejemplo, el sujeto tiene al menos 40 años.
- 10 **[0112]** Los sujetos con antecedentes familiares de cáncer colorrectal pueden someterse antes a un cribado. Por ejemplo, estos sujetos pueden ser examinados a partir de al menos 30, al menos 31, al menos 32, al menos 33, al menos 34, al menos 35, al menos 36, al menos 37 años de edad o más.
- 15 **[0113]** En otro ejemplo, los sujetos evaluados utilizando los métodos de la invención han tenido una prueba de sangre oculta en heces positiva. En otros ejemplos, los sujetos tienen antecedentes personales de pólipos adenomatosos o antecedentes personales de enfermedad inflamatoria intestinal (colitis ulcerosa o enfermedad de Crohn).
- 20 **[0114]** En otro ejemplo, los métodos de la invención pueden usarse para evaluar el riesgo de un sujeto humano de desarrollar cáncer colorrectal con síntomas que pueden ser indicativos de cáncer colorrectal. En el contexto del cáncer colorrectal, el método sería aplicable a un sujeto con una prueba de cribado fecal oculta positiva o a un sujeto que acuda a la clínica con síntomas como cambio en los hábitos intestinales, incluyendo diarrea o estreñimiento, cambio en la consistencia de las heces, hemorragia rectal, molestias abdominales persistentes, como calambres, defecación incompleta, gases o dolor.
- 25 **[0115]** Los métodos de la invención pueden utilizarse para evaluar el riesgo en sujetos masculinos y femeninos. Sin embargo, en un ejemplo, el sujeto es varón.
- 30 **[0116]** Los métodos de la invención pueden utilizarse para evaluar el riesgo de desarrollar cáncer colorrectal en sujetos humanos de diversos orígenes étnicos. Es bien sabido que a lo largo del tiempo se han ido mezclando diferentes orígenes étnicos. Aunque en la práctica esto no influye en la capacidad de un experto para practicar los métodos aquí descritos, puede ser conveniente identificar el origen étnico del sujeto. En este caso, la etnia del sujeto humano puede ser declarada por él mismo. Por ejemplo, se puede pedir a los sujetos que identifiquen su origen étnico en respuesta a esta pregunta: "¿A qué grupo étnico pertenece?" En otro ejemplo, la etnia del sujeto puede derivarse de los historiales médicos tras obtener el consentimiento correspondiente del sujeto o de la opinión u observaciones de un clínico.
- 35 **[0117]** En un ejemplo, el sujeto puede clasificarse como *Caucasoide*, *Australoide*, *Mongoloide* y *Negroide* basándose en la antropología física. En algunas realizaciones, el sujeto puede ser caucásico, afroamericano, hispano, asiático, indio o latino. En un ejemplo, el sujeto es caucásico. Por ejemplo, el sujeto puede ser europeo.
- 40 **[0118]** Un sujeto de origen predominantemente europeo, ya sea directo o indirecto a través de la ascendencia, con piel blanca se considera caucásico en el contexto de la presente divulgación. Un caucásico puede tener, por ejemplo, al menos un 75% de ascendencia caucásica (por ejemplo, pero sin limitarse a ello, el sujeto tiene al menos tres abuelos caucásicos).
- 45 **[0119]** Un sujeto americano con ascendencia predominantemente africana central o meridional, por ejemplo, al menos un 75% de ascendencia africana central o meridional, y piel negra se considera afroamericano en el contexto de la presente divulgación. Un principio similar se aplica, por ejemplo, a los súbditos de ascendencia centroafricana o sudafricana que viven en otros países (por ejemplo, Gran Bretaña, Canadá u Holanda).
- 50 **[0120]** Un sujeto predominantemente originario de España o de un país hispanohablante, como un país de América Central o del Sur, ya sea directa o indirectamente por ascendencia, se considera hispano en el contexto de la presente divulgación. Un sujeto hispano puede tener, por ejemplo, al menos un 75% de ascendencia hispana.
- Cribado Rutinario
- 55 **[0121]** Las pruebas de sangre oculta en heces y la colonoscopia/ sigmoidoscopia reducen la mortalidad por cáncer colorrectal, pero son caras de ofrecer de forma rutinaria a un gran número de sujetos. En consecuencia, es conveniente identificar a la población adecuada para realizar el cribado. En un ejemplo, los métodos de la invención pueden utilizarse para determinar la necesidad de realizar pruebas diagnósticas rutinarias de cáncer colorrectal en un sujeto humano. Dicho cribado rutinario puede incluir pruebas de sangre oculta en heces o colonoscopia/sigmoidoscopia a intervalos de tiempo predeterminados como los comentados anteriormente.
- 60 **[0122]** En un ejemplo, la necesidad de pruebas diagnósticas rutinarias de un sujeto humano para cáncer colorrectal se determina en base al número de alelos de riesgo detectados. Un experto en la materia apreciaría que cada uno de los polimorfismos de nucleótido único puede estar presente hasta dos veces en el genoma diploide somático del sujeto. Así, por ejemplo, una evaluación de 28 polimorfismos de un solo nucleótido puede dar lugar a la detección de 56 alelos. En
- 65

otro ejemplo, una evaluación de 45 polimorfismos de un solo nucleótido puede dar como resultado la detección de 90 alelos. Una parte de los alelos detectados pueden ser alelos de riesgo. El número de alelos de riesgo detectados es relevante para el riesgo del sujeto de desarrollar un cáncer de colon.

5 **[0123]** En un ejemplo, cuando se tiene en cuenta que cada uno de los polimorfismos de nucleótido único puede estar presente hasta dos veces en el genoma diploide somático del sujeto, un sujeto que tiene al menos 41, al menos 42, al menos 43, al menos 44, al menos 45, al menos 46, al menos 47, al menos 48, al menos 49 al menos 50, al menos 51, al menos 52, al menos 53, al menos 54, al menos 55, al menos 56, al menos 57, al menos 58, al menos 59, al menos 60 o más alelos de riesgo de los polimorfismos de nucleótido único debe inscribirse en un programa de cribado oculto en heces, 10 cribado colonoscópico o sigmoidoscópico. Por ejemplo, los sujetos con al menos 44 alelos de riesgo de los polimorfismos de un solo nucleótido deben inscribirse en un programa de cribado fecal oculto, colonoscópico o sigmoidoscópico. En un ejemplo, los sujetos de al menos 49 años de edad con al menos 44 alelos de riesgo de los polimorfismos de nucleótido único deben inscribirse en un programa de cribado colonoscópico o sigmoidoscópico.

15 **[0124]** En otro ejemplo, los sujetos con al menos 46 alelos de riesgo de los polimorfismos de nucleótido único deben inscribirse en un programa de cribado fecal oculto, colonoscópico o sigmoidoscópico. En este ejemplo, los sujetos de al menos 47 años de edad con al menos 46 alelos de riesgo de los polimorfismos de nucleótido único deben inscribirse en un programa de cribado colonoscópico o sigmoidoscópico.

20 **[0125]** En otro ejemplo, la necesidad de pruebas diagnósticas rutinarias de cáncer colorrectal en un sujeto humano se determina basándose en la clasificación de riesgo del sujeto dentro de una población de sujetos. Por ejemplo, si la evaluación sitúa al sujeto en el top 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1% de sujetos de una población con riesgo de desarrollar cáncer colorrectal, entonces el sujeto se inscribe en un programa de cribado oculto en heces, colonoscópico o sigmoidoscópico.

25 **[0126]** En un ejemplo, el riesgo genético se calcula en función de: $SNP_1 \times SNP_2 \times SNP_3 \times SNP_4 \times SNP_5 \times SNP_6 \times SNP_7 \times SNP_x$ y los sujetos con un riesgo superior a aproximadamente el 5,9% se inscriben en un programa de cribado oculto en heces, colonoscópico o sigmoidoscópico. En otro ejemplo, los sujetos con un riesgo superior a aproximadamente 6,0, 6,1, 6,2, 6,3, 6,4, 6,5, 6,6, 6,7, 6,8, 6,9, 7,0, 7,1, 7,2, 7,3, 7,4% o más se inscriben en un programa de cribado oculto en 30 heces, colonoscópico o sigmoidoscópico.

[0127] En otro ejemplo, el riesgo combinado (es decir, riesgo clínico x riesgo genético) se calcula a partir de: [riesgo clínico asociado a tener un pariente de primer grado con cáncer colorrectal] x $SNP_1 \times SNP_2 \times SNP_3 \times SNP_4 \times SNP_5 \times SNP_6 \times SNP_7 \times SNP_x$ y los sujetos que tienen un riesgo superior a aproximadamente el 11,5% se inscriben en un programa de 35 cribado oculto en heces, colonoscópico o sigmoidoscópico. En otro ejemplo, los sujetos que tienen un riesgo superior a aproximadamente 12, 12,5, 13, 13,1, 13,2, 13,3, 13,4, 13,5, 14% o más se inscriben en un programa de cribado fecal oculto, colonoscópico o sigmoidoscópico.

40 **[0128]** En otro ejemplo, los métodos de la invención se incorporan a un método de cribado del cáncer colorrectal en un sujeto. En este ejemplo, el riesgo de un sujeto de desarrollar cáncer colorrectal se evalúa utilizando los métodos de la invención y el sujeto se somete a un cribado rutinario de cáncer colorrectal mediante colonoscopia o sigmoidoscopia si se evalúa que tiene riesgo de desarrollar cáncer colorrectal.

45 **[0129]** Los métodos de la invención también pueden utilizarse en combinación con otros métodos o "prueba(s) adicional(es)" para proporcionar una evaluación del riesgo de desarrollar cáncer colorrectal. En este ejemplo, los resultados de múltiples pruebas pueden ayudar a un clínico a determinar si es necesaria una prueba más definitiva, como una colonoscopia o una sigmoidoscopia. En un ejemplo, los métodos de la invención pueden realizarse en combinación con una prueba de sangre oculta en heces.

50 Rendimiento del Método

[0130] En diversas realizaciones, el rendimiento del método se caracteriza por un área bajo la curva (AUC) de al menos aproximadamente 0,61, al menos aproximadamente 0,62, al menos aproximadamente 0,63.

55 **[0131]** En diversas realizaciones, la sensibilidad alcanzada por los métodos de la presente divulgación es de al menos aproximadamente 50%, al menos aproximadamente 60%, al menos aproximadamente 70%, al menos aproximadamente 71%, al menos aproximadamente 72%, al menos aproximadamente 73%, al menos aproximadamente 74%, al menos aproximadamente 75%, al menos aproximadamente 76%, al menos aproximadamente 77%, al menos aproximadamente 78%, al menos aproximadamente del 79%, al menos aproximadamente del 80%, al menos aproximadamente del 81%, al menos aproximadamente del 82%, al menos aproximadamente del 83%, al menos aproximadamente del 84%, al menos aproximadamente del 85%, al menos aproximadamente del 86%, al menos aproximadamente del 87%, al menos aproximadamente del 88%, al menos aproximadamente del 89%, al menos aproximadamente del 90%, al menos aproximadamente del 91%, al menos aproximadamente del 92%, al menos aproximadamente del 93%, al menos aproximadamente del 94%, al menos aproximadamente del 95%.

65 **[0132]** En diversas realizaciones, la especificidad lograda por los métodos de la presente divulgación es de al menos

aproximadamente 50%, al menos aproximadamente 60%, al menos aproximadamente 70%, al menos aproximadamente 71%, al menos aproximadamente 72%, al menos aproximadamente 73%, al menos aproximadamente 74%, al menos aproximadamente 75%, al menos aproximadamente 76%, al menos aproximadamente 77%, al menos aproximadamente 78%, al menos aproximadamente del 79%, al menos aproximadamente del 80%, al menos aproximadamente del 81%, al menos aproximadamente del 82%, al menos aproximadamente del 83%, al menos aproximadamente del 84%, al menos aproximadamente del 85%, al menos aproximadamente del 86%, al menos aproximadamente del 87%, al menos aproximadamente del 88%, al menos aproximadamente del 89%, al menos aproximadamente del 90%, al menos aproximadamente del 91%, al menos aproximadamente del 92%, al menos aproximadamente del 93%, al menos aproximadamente del 94%, al menos aproximadamente del 95%.

Tratamiento

[0133] Una alta propensión genética al cáncer colorrectal puede tratarse como una advertencia para iniciar un tratamiento profiláctico o terapéutico. Así, tras realizar los métodos de la invención, se puede prescribir o administrar un tratamiento al sujeto. En algunas realizaciones, los métodos de la invención se refieren a una terapia contra el cáncer colorrectal para su uso en la prevención o reducción del riesgo de cáncer colorrectal en un sujeto humano en riesgo del mismo. Como también se describe en el presente documento, al sujeto se le puede prescribir o administrar un agente terapéutico o profiláctico. Por ejemplo, al sujeto se le puede prescribir o administrar un quimiopreventivo. En otros ejemplos, al sujeto se le pueden prescribir o administrar antiinflamatorios no esteroideos como aspirina, buprofeno, paracetamol y naproxeno o terapia hormonal (estrógeno más progestina). En otro ejemplo, el tratamiento puede incluir una intervención conductual, como la manipulación de la dieta de los sujetos. Las modificaciones dietéticas ejemplares incluyen el aumento de fibra, ácidos grasos monoinsaturados y/o aceite de pescado.

Preparación y Análisis de Muestras

[0134] En la realización de los métodos de la invención, se requiere una muestra biológica de un sujeto. Se considera que términos como "muestra" y "espécimen" son términos que, en su contexto, pueden utilizarse indistintamente en la presente divulgación. Se puede utilizar cualquier material biológico como la muestra mencionada, siempre y cuando se pueda derivar del sujeto y se pueda aislar y analizar el ADN de acuerdo con los métodos de la presente divulgación. Las muestras suelen tomarse, previo consentimiento informado, de un paciente mediante métodos de laboratorio médico estándar. La muestra puede estar en una forma tomada directamente del paciente, o puede estar al menos parcialmente procesada (purificada) para eliminar al menos parte del material no ácido nucleico.

[0135] Las "muestras biológicas" ejemplares incluyen fluidos corporales (sangre, saliva, orina, etc.), biopsia, tejido y/o desechos del paciente. Así, biopsias de tejido, heces, esputo, saliva, sangre, linfa, lágrimas, sudor, orina, secreciones vaginales, o similares pueden ser fácilmente analizados para SNP, al igual que cualquier tejido de interés que contenga los ácidos nucleicos apropiados. En una realización, la muestra biológica es una muestra de células de la mejilla.

[0136] En algunas realizaciones, la muestra es una muestra de sangre. Una muestra de sangre puede tratarse para eliminar células concretas mediante diversos métodos, como la centrifugación, la cromatografía de afinidad (por ejemplo, medios inmunoabsorbentes), la inmunoselección y la filtración, si es necesario. Así, en un ejemplo, la muestra puede comprender un tipo celular específico o una mezcla de tipos celulares aislados directamente del sujeto o purificados a partir de una muestra obtenida del sujeto. En un ejemplo, la muestra biológica son células mononucleares de sangre periférica (pBMC). Se conocen varios métodos para purificar subpoblaciones de células. Por ejemplo, las pBMC pueden purificarse a partir de sangre total utilizando diversos métodos conocidos de centrifugación basados en Ficoll (por ejemplo, centrifugación en gradiente de densidad Ficoll-Hypaque).

[0137] Se puede extraer ADN de la muestra para detectar SNP. En un ejemplo, el ADN es ADN genómico. Los expertos en la materia conocen varios métodos para aislar el ADN, en particular el ADN genómico. En general, los métodos conocidos implican la disrupción y lisis del material de partida, seguida de la eliminación de proteínas y otros contaminantes y, por último, la recuperación del ADN. Por ejemplo, las técnicas de precipitación con alcohol, extracción orgánica con fenol/cloroformo y salazón se han utilizado durante muchos años para extraer y aislar el ADN. Existen varios kits comerciales para la extracción de ADN genómico (Qiagen, Life technologies; Sigma). La pureza y la concentración del ADN pueden evaluarse mediante diversos métodos, por ejemplo, la espectrofotometría.

Estrategias de detección de marcadores

[0138] En los métodos de la invención pueden utilizarse cebadores de amplificación para amplificar marcadores (p. ej., loci marcadores) y sondas adecuadas para detectar dichos marcadores o para genotipar una muestra con respecto a múltiples alelos marcadores. Por ejemplo, la selección de cebadores para la PCR de largo alcance se describe en los documentos US 10/042,406 y US 10/236,480; para la PCR de corto alcance, el documento US 10/341,832 proporciona orientación con respecto a la selección de cebadores. Además, existen programas públicos como "Oligo" para el diseño de cebadores. Con el software de selección y diseño de cebadores disponible, la secuencia del genoma humano disponible públicamente y las localizaciones de los polimorfismos, un experto en la materia puede construir cebadores para amplificar los SNP para practicar el método de la invención. Además, se apreciará que la sonda precisa a utilizar para la detección de un ácido nucleico que comprende un SNP (p. ej., un amplicón que comprende el SNP) puede variar,

p. ej., cualquier sonda que pueda identificar la región de un amplicón marcador a detectar puede utilizarse junto con la presente divulgación. Además, la configuración de las sondas de detección puede, por supuesto, variar.

[0139] En la Tabla 3 se proporcionan ejemplos de cebadores oligonucleotídicos útiles para amplificar ácidos nucleicos que comprenden SNP que se sabe que están asociados con un cáncer colorrectal. Como apreciará el experto, la secuencia de la región genómica a la que hibridan estos oligonucleótidos puede utilizarse para diseñar cebadores que sean más largos en el extremo 5' y/o 3', posiblemente más cortos en el extremo 5' y/o 3' (siempre que la versión truncada pueda seguir utilizándose para la amplificación) o que no compartan ninguna similitud de secuencia. (siempre que la versión truncada pueda seguir utilizándose para la amplificación), que tengan una o unas pocas diferencias de nucleótidos (pero que, no obstante, puedan seguir utilizándose para la amplificación), o que no compartan ninguna similitud de secuencia con los proporcionados, pero que estén diseñados basándose en secuencias genómicas cercanas a las que hibridan los oligonucleótidos específicamente proporcionados y que puedan seguir utilizándose para la amplificación.

Tabla 3. Cebadores y sondas TaqMan para los seis SNP de mayor riesgo mostrados en la Tabla 1.

SNP	Cebador directo	Cebador inverso
rs72647484	TGCAGCAAGTGCTGAGAAG (SEQ ID NO:1)	CCCATTTGTTACCAGTATGAAGAGT (SEQ ID NO:2)
rs3987	AGACACTCTCCTCTGTTGATT (SEQ ID NO:3)	GGACATCAAATAATGTGCCTAGAA (SEQ ID NO:4)
rs35509282	CCTGAGTAGCTGGGACTACA (SEQ ID NO:5)	TCGAGACCATCCTGGCTAA (SEQ ID NO:6)
rs16892766	AACGGTCAGACGCAAACA (SEQ ID NO:7)	GACGGCAATAAATCTTCCATGAG (SEQ ID NO:8)
rs6983267	CCTTTGAGCTCAGCAGATGAA (SEQ ID NO:9)	GGGTTCTGCCCCTTTGATT (SEQ ID NO:10)
rs744166	TTGGGCCACACAGTCTCTAA (SEQ ID NO:11)	TGAGTTGCTGTGGCTGTAATG (SEQ ID NO:12)

[0140] Los cebadores pueden ser radiomarcados, o marcados por cualquier medio adecuado (p. ej., utilizando una etiqueta fluorescente no radiactiva), para permitir la visualización rápida de amplicóns de diferentes tamaños después de una reacción de amplificación sin ninguna etapa adicional de marcación o visualización. Alternativamente, los cebadores pueden no estar marcados, y los amplicóns pueden visualizarse tras su resolución de tamaño, p. ej., tras electroforesis en gel de agarosa o acrilamida. Por ejemplo, la tinción con bromuro de etidio de los amplicóns de la PCR tras la resolución del tamaño permite visualizar los amplicóns de distinto tamaño.

[0141] No se pretende que los cebadores se limiten a generar un amplicón de un tamaño determinado. Por ejemplo, los cebadores utilizados para amplificar los loci y alelos marcadores pueden no limitarse a amplificar toda la región del locus pertinente, o cualquier subregión del mismo. Los cebadores pueden generar un amplicón de cualquier longitud adecuada para la detección. Por ejemplo, la amplificación del marcador puede producir un amplicón de al menos 20 nucleótidos de longitud, o alternativamente, de al menos 50 nucleótidos de longitud, o alternativamente, de al menos 100 nucleótidos de longitud, o alternativamente, de al menos 200 nucleótidos de longitud. Pueden detectarse amplicóns de cualquier tamaño utilizando las diversas tecnologías aquí descritas. Las diferencias en la composición o el tamaño de las bases pueden detectarse mediante métodos convencionales como la electroforesis.

[0142] De hecho, se apreciará que la amplificación no es un requisito para la detección de marcadores, por ejemplo uno puede detectar directamente ADN genómico no amplificado simplemente realizando un Southern blot en una muestra de ADN genómico.

[0143] Típicamente, los marcadores moleculares pueden ser detectados por cualquier método establecido disponible en el arte, incluyendo, sin limitación, hibridación específica de alelo (ASH), detección de extensión de nucleótido único, hibridación de matriz (opcionalmente incluyendo ASH), u otros métodos para detectar polimorfismos de nucleótido único, detección de polimorfismo de longitud de fragmento amplificado (AFLP), detección de secuencia variable amplificada, detección de ADN polimórfico amplificado aleatoriamente (RAPD), detección de polimorfismo de longitud de fragmento de restricción (RFLP), detección de replicación de secuencia autosostenida, detección de repetición de secuencia simple (SSR) y detección de polimorfismos de conformación de cadena simple (SSCP).

[0144] Algunas técnicas para detectar marcadores genéticos utilizan la hibridación de un ácido nucleico de sonda a ácidos nucleicos correspondientes al marcador genético (p. ej., ácidos nucleicos amplificados producidos utilizando ADN genómico como molde). Los formatos de hibridación, que incluyen, entre otros: ensayos de hibridación en fase de solución, en fase sólida, en fase mixta o in situ, son útiles para la detección de alelos. Una extensa guía sobre la hibridación de

ácidos nucleicos se encuentra en Tijssen (1993) y Sambrook et al. (supra).

[0145] También puede realizarse la detección por PCR utilizando sondas de oligonucleótidos fluorogénicos de doble marcación, comúnmente denominadas sondas "TaqMan™". Estas sondas se componen de oligodesoxinucleótidos cortos (p. ej., 20-25 bases) marcados con dos colorantes fluorescentes diferentes. En el extremo 5' de cada sonda se encuentra un colorante informador, y en el extremo 3' de cada sonda se encuentra un colorante amortiguador. La secuencia de la sonda oligonucleotídica es complementaria a una secuencia diana interna presente en un amplicón PCR. Cuando la sonda está intacta, se produce una transferencia de energía entre los dos fluoróforos y la emisión del indicador es apagada por el inhibidor mediante FRET. Durante la fase de extensión de la PCR, la sonda es escindida por la actividad 5' nucleasa de la polimerasa utilizada en la reacción, liberando así el indicador del inhibidor del oligonucleótido y produciendo un aumento de la intensidad de emisión del indicador. En consecuencia, las sondas TaqMan™ son oligonucleótidos que tienen un marcador y un inhibidor, donde el marcador se libera durante la amplificación por la acción exonucleasa de la polimerasa utilizada en la amplificación. Esto proporciona una medida en tiempo real de la amplificación durante la síntesis. Existe una gran variedad de reactivos TaqMan™ disponibles en el mercado, p. ej., en Applied Biosystems (sede de la división en Foster City, California), así como en diversos proveedores especializados como Biosearch Technologies (p. ej., sondas de inhibidor de agujero negro). En el documento WO 92/02638, p. ej., se ofrecen más detalles sobre las estrategias de sonda de doble marcación.

[0146] Otros métodos similares incluyen, p. ej., la transferencia de energía por resonancia de fluorescencia entre dos sondas hibridadas adyacentemente, p. ej., utilizando el formato "LightCycler®" descrito en US 6,174,670.

[0147] La detección basada en matrices puede realizarse utilizando matrices disponibles comercialmente, p. ej., de Affymetrix (Santa Clara, Calif.) u otros fabricantes. Entre las revisiones sobre el funcionamiento de las matrices de ácidos nucleicos se incluyen Sapolsky et al. (1999); Lockhart (1998); Fodor (1997a); Fodor (1997b) y Chee et al. (1996). La detección basada en matrices es uno de los métodos preferidos para la identificación de marcadores en muestras, debido a la naturaleza inherente de alto rendimiento de la detección basada en matrices.

[0148] La muestra de ácido nucleico a analizar es aislada, amplificada y, típicamente, marcada con biotina y/o un grupo indicador fluorescente. A continuación, la muestra de ácido nucleico marcada se incuba con la matriz utilizando una estación de fluidica y un horno de hibridación. La matriz puede lavarse y teñirse o contrateñirse, según convenga al método de detección. Tras la hibridación, el lavado y la tinción, la matriz se introduce en un escáner, donde se detectan los patrones de hibridación. Los datos de hibridación se recogen como luz emitida por los grupos informadores fluorescentes ya incorporados al ácido nucleico marcado, que ahora está unido a la matriz de sondas. Las sondas que coinciden más claramente con el ácido nucleico marcado producen señales más fuertes que las que tienen coincidencias erróneas. Dado que se conocen la secuencia y la posición de cada sonda en la matriz, por complementariedad se puede identificar la identidad de la muestra de ácido nucleico aplicada a la matriz de sondas.

Correlación de Marcadores con el Riesgo de Cáncer

[0149] Las correlaciones entre los SNP y el riesgo de cáncer colorrectal pueden realizarse mediante cualquier método que pueda identificar una relación entre un alelo y un mayor riesgo de cáncer, o una combinación de alelos y un mayor riesgo de cáncer. Por ejemplo, los alelos en genes o loci definidos en el presente documento pueden correlacionarse con un mayor riesgo de cáncer colorrectal. Por lo general, estos métodos consisten en consultar una tabla de consulta que contiene las correlaciones entre los alelos del polimorfismo y el riesgo de cáncer. La tabla puede incluir datos de múltiples relaciones alelo-riesgo y puede tener en cuenta los efectos aditivos u otros efectos de orden superior de múltiples relaciones alelo-riesgo, p. ej., mediante el uso de herramientas estadísticas como el análisis de componentes principales, algoritmos heurísticos, etc.

[0150] La correlación de un marcador con un riesgo de cáncer incluye opcionalmente la realización de una o más pruebas estadísticas de correlación. Se conocen muchas pruebas estadísticas y la mayoría se implementan por ordenador para facilitar el análisis. Se conocen diversos métodos estadísticos para determinar asociaciones/correlaciones entre rasgos fenotípicos y marcadores biológicos, que pueden aplicarse a la presente divulgación. Hartl (1981). En Lynch y Walsh (1998) se describen diversos modelos estadísticos apropiados. Estos modelos pueden, por ejemplo, proporcionar correlaciones entre valores genotípicos y fenotípicos, caracterizar la influencia de un locus en el riesgo de cáncer, ordenar la relación entre el entorno y el genotipo, determinar la dominancia o penetrancia de los genes, determinar los efectos epigenéticos maternos y de otro tipo, determinar los componentes principales en un análisis (mediante el análisis de componentes principales, o "ACP"), y similares. Las referencias citadas en estos textos proporcionan muchos más detalles sobre los modelos estadísticos para correlacionar los marcadores y el riesgo de cáncer.

[0151] Además de los métodos estadísticos estándar para determinar la correlación, pueden utilizarse otros métodos que determinan correlaciones mediante el reconocimiento de patrones y el entrenamiento, como el uso de algoritmos genéticos, para determinar correlaciones entre marcadores y riesgo de cáncer. Esto resulta especialmente útil a la hora de identificar correlaciones de orden superior entre múltiples alelos y el riesgo de cáncer. A modo de ejemplo, los enfoques de redes neuronales pueden acoplarse a la programación de tipo algoritmo genético para el desarrollo heurístico de un modelo de espacio de datos estructura-función que determine las correlaciones entre la información genética y los resultados fenotípicos.

[0152] En cualquier caso, esencialmente cualquier prueba estadística puede aplicarse en un modelo implementado por ordenador, por métodos de programación estándar, o utilizando cualquiera de una variedad de paquetes de software "off the shelf" que realizan tales análisis estadísticos, incluyendo, p. ej., los señalados anteriormente y los que están disponibles comercialmente, p. ej., de Partek Incorporated (St. Peters, Mo.; www.partek.com), p. ej., que proporcionan software para el reconocimiento de patrones (p. ej., que proporcionan Partek Pro 2000 Pattern Recognition Software).

[0153] Se pueden encontrar detalles adicionales sobre los estudios de asociación en los documentos US 10/106,097, US 10/042,819, US 10/286,417, US 10/768,788, US 10/447,685, US 10/970,761 y US 7,127,355.

[0154] Los sistemas para realizar las correlaciones anteriores son también una característica de la divulgación. Típicamente, el sistema incluirá instrucciones que correlacionan la presencia o ausencia de un alelo (ya sea detectado directamente o, p. ej., a través de los niveles de expresión) con un riesgo de cáncer previsto.

[0155] Opcionalmente, las instrucciones del sistema también pueden incluir software que acepte información de diagnóstico asociada con cualquier información alélica detectada, p. ej., un diagnóstico de que un sujeto con el alelo relevante tiene un riesgo de cáncer particular. Este software puede ser de naturaleza heurística, utilizando las asociaciones introducidas para mejorar la precisión de las tablas de consulta y/o la interpretación de las tablas de consulta por parte del sistema. Más arriba se describen diversos enfoques de este tipo, como las redes neuronales, la modelización de Markov y otros análisis estadísticos.

Elaboración de Perfiles Polimórficos

[0156] En el presente documento se describen métodos para determinar el perfil polimórfico de un individuo en los SNP descritos en el presente documento (Tabla 6) o SNP en desequilibrio de ligamiento con uno o más de ellos.

[0157] El perfil polimórfico constituye las formas polimórficas que ocupan los diversos sitios polimórficos en un individuo. En un genoma diploide, dos formas polimórficas, iguales o diferentes entre sí, suelen ocupar cada sitio polimórfico. Así, el perfil polimórfico en los sitios X e Y puede representarse en la forma X (x1, x1), e Y (y1, y2), en la que x1, x1 representa dos copias del alelo x1 que ocupan el sitio X y y1, y2 representan alelos heterocigotos que ocupan el sitio Y.

[0158] El perfil polimórfico de un individuo puede ser puntuado por comparación con las formas polimórficas asociadas con la susceptibilidad al cáncer colorrectal que ocurren en cada sitio. La comparación puede realizarse en al menos, p. ej., 1, 2, 5, 10, 25, 50, o todos los sitios polimórficos, y opcionalmente, otros en desequilibrio de ligamiento con ellos. Los sitios polimórficos pueden analizarse en combinación con otros sitios polimórficos.

[0159] El perfil polimórfico es útil, por ejemplo, en la selección de agentes para afectar el tratamiento o la profilaxis del cáncer colorrectal en un individuo dado. Es probable que los individuos con perfiles polimórficos similares respondan a los agentes de forma parecida.

Método informatizado

[0160] Los métodos de la invención pueden ser implementados por un sistema como un método implementado por ordenador. Por ejemplo, el sistema puede ser un sistema informático que comprenda uno o una pluralidad de procesadores que puedan funcionar juntos (denominados por comodidad "procesador") conectados a una memoria. La memoria puede ser un medio no transitorio legible por ordenador, como un disco duro, un disco de estado sólido o un CD-ROM. El software, es decir, las instrucciones ejecutables o el código de programa, como el código de programa agrupado en módulos de código, puede almacenarse en la memoria, y puede, cuando es ejecutado por el procesador, hacer que el sistema informático realice funciones tales como determinar que debe realizarse una tarea para ayudar a un usuario a determinar el riesgo de un sujeto humano de desarrollar cáncer colorrectal recibiendo datos que indican el riesgo genético y, opcionalmente, el riesgo clínico de que el sujeto desarrolle cáncer colorrectal, en los que el riesgo genético se ha obtenido detectando, en una muestra biológica procedente del sujeto, la presencia de al menos 45 polimorfismos de nucleótido único que figuran en la Tabla 1 ; procesar los datos para obtener el riesgo de un sujeto humano de desarrollar cáncer colorrectal; emitir la presencia del riesgo de un sujeto humano de desarrollar cáncer colorrectal.

[0161] Por ejemplo, la memoria puede comprender código de programa que cuando es ejecutado por el procesador hace que el sistema determine la presencia de los al menos 45 polimorfismos de nucleótido único mostrados en la Tabla 1, o reciba datos que indiquen la presencia de los al menos 45 polimorfismos de nucleótido único seleccionados de la Tabla 1; procese los datos para obtener el riesgo de un sujeto humano de desarrollar cáncer colorrectal; informe del riesgo de un sujeto humano de desarrollar cáncer colorrectal. Así, en una realización, el código de programa hace que el sistema determine el "riesgo genético".

[0162] En otra realización, la memoria puede comprender un código de programa que, cuando es ejecutado por el procesador, hace que el sistema determine la presencia de al menos 45 polimorfismos de nucleótido único mostrados en la Tabla 1, o reciba datos que indiquen la presencia de al menos 45 polimorfismos de nucleótido único mostrados en la Tabla 1, y reciba o determine datos de riesgo clínico para el sujeto; procese los datos para combinar los datos de riesgo

genético con los datos de riesgo clínico para obtener el riesgo del sujeto de desarrollar cáncer colorrectal; informe del riesgo de un sujeto humano de desarrollar cáncer colorrectal. Por ejemplo, el código del programa puede hacer que el sistema combine datos de evaluación del riesgo clínico × riesgo genético.

5 **[0163]** En algunas realizaciones, el sistema puede estar acoplado a una interfaz de usuario para permitir al sistema recibir información de un usuario y/o emitir o mostrar información. Por ejemplo, la interfaz de usuario puede comprender una interfaz gráfica de usuario, una interfaz de usuario de voz o una pantalla táctil. En un ejemplo, la interfaz de usuario es una plataforma de matriz SNP.

10 **[0164]** En algunas realizaciones, el sistema puede estar configurado para comunicarse con al menos un dispositivo o servidor remoto a través de una red de comunicaciones tal como una red de comunicaciones inalámbrica. Por ejemplo, el sistema puede estar configurado para recibir información del dispositivo o servidor a través de la red de comunicaciones y para transmitir información al mismo dispositivo o servidor o a otro diferente a través de la red de comunicaciones. En otras realizaciones, el sistema puede estar aislado de la interacción directa del usuario.

15 **[0165]** En algunas realizaciones, la ejecución de los métodos de la invención para evaluar el riesgo de un sujeto de desarrollar cáncer colorrectal permite establecer una regla de diagnóstico o pronóstico basada en el riesgo genético de que el sujeto desarrolle cáncer colorrectal. Por ejemplo, la regla de diagnóstico o pronóstico puede basarse en el riesgo genético relativo a un nivel de riesgo de control, estándar o umbral. En otro ejemplo, la regla diagnóstica o pronóstica puede basarse en el riesgo genético y clínico combinado en relación con un nivel de riesgo de control, estándar o umbral.

20 **[0166]** En algunas realizaciones, la regla de diagnóstico o pronóstico se basa en la aplicación de un algoritmo estadístico y de aprendizaje automático. Un algoritmo de este tipo utiliza relaciones entre una población de SNP y el estado de la enfermedad observado en los datos de entrenamiento (con estado de enfermedad conocido) para inferir relaciones que luego se utilizan para determinar el riesgo de un sujeto humano de desarrollar cáncer colorrectal en sujetos con un riesgo desconocido. Se emplea un algoritmo que proporciona el riesgo de que un sujeto humano desarrolle cáncer colorrectal. El algoritmo realiza una función de análisis multivariante o univariante.

Kits y Productos

30 **[0167]** También se divulga en el presente documento, pero no forma parte de la invención reivindicada, un kit que comprende al menos 28 conjuntos de cebadores para amplificar 28 o más ácidos nucleicos, en el que los 28 o más ácidos nucleicos comprenden un polimorfismo de nucleótido único seleccionado de la Tabla 1, o un polimorfismo de nucleótido único en desequilibrio de ligamiento con uno o más de ellos.

35 **[0168]** El kit puede comprender por lo menos 28, por lo menos 29, por lo menos 30, por lo menos 31, por lo menos 32, por lo menos 33, por lo menos 34, por lo menos 35, por lo menos 36, por lo menos 37, por lo menos 38, por lo menos 39, por lo menos 40, por lo menos 41, por lo menos 42, por lo menos 43 al menos 44, al menos 45 conjuntos de cebadores para amplificar ácidos nucleicos que comprenden un polimorfismo de nucleótido único seleccionado de la Tabla 1, o un polimorfismo de nucleótido único en desequilibrio de ligamiento con uno o más de ellos.

40 **[0169]** Como pueden apreciar los expertos en la materia, una vez identificado un SNP, se pueden diseñar cebadores para amplificar el SNP de forma rutinaria. Existen varios programas informáticos gratuitos que pueden sugerir cebadores adecuados para amplificar los SNP de interés.

45 **[0170]** Una vez más, sería conocido por los expertos en la materia que los cebadores PCR de un par de cebadores PCR pueden diseñarse para amplificar específicamente una región de interés a partir de ADN humano. En el contexto de la presente divulgación, la región de interés contiene la variación de base única (por ejemplo, polimorfismo de un solo nucleótido, SNP) que se genotipará. Cada cebador PCR de un par de cebadores PCR puede colocarse adyacente a una variación monobásica concreta en sitios opuestos de la variación de la secuencia de ADN. Además, los cebadores PCR pueden diseñarse para evitar cualquier variación de secuencia de ADN conocida y secuencias de ADN repetitivas en sus sitios de unión al cebador PCR.

50 **[0171]** El kit puede comprender además otros reactivos necesarios para realizar una reacción de amplificación, como un tampón, nucleótidos y/o una polimerasa, así como reactivos para extraer ácidos nucleicos de una muestra.

55 **[0172]** La detección basada en matrices es un método preferido para evaluar los SNP de la divulgación en muestras, debido a la naturaleza inherente de alto rendimiento de la detección basada en matrices. En la bibliografía se han descrito diversas matrices de sondas que pueden utilizarse en el contexto de la presente divulgación para la detección de SNP que puedan correlacionarse con el cáncer colorrectal. Por ejemplo, en una realización de la divulgación se utilizan chips de matrices de sondas de ADN. El reconocimiento del ADN de la muestra por el conjunto de sondas de ADN tiene lugar mediante la hibridación del ADN. Cuando una muestra de ADN se hibrida con una matriz de sondas de ADN, la muestra se une a aquellas sondas que son complementarias a la secuencia de ADN de la muestra. Evaluando a qué sondas hibrida más fuertemente el ADN de muestra de un individuo, es posible determinar si una secuencia conocida de ácido nucleico está presente o no en la muestra, determinando así si un marcador encontrado en el ácido nucleico está presente.

[0173] Así pues, también se divulga en el presente documento pero no forma parte de la invención reivindicada un conjunto genético que comprende al menos 28 conjuntos de sondas para hibridar con 28 o más ácidos nucleicos, en el que los 28 o más ácidos nucleicos comprenden un polimorfismo de nucleótido único seleccionado de la Tabla 1, o un polimorfismo de nucleótido único en desequilibrio de ligamiento con uno o más de ellos. En una realización, la matriz comprende al menos 28, al menos 29, al menos 30, al menos 31, al menos 32, al menos 33, al menos 34, al menos 35, al menos 36, al menos 37, al menos 38, al menos 39, al menos 40, al menos 41, al menos 42, al menos 43, al menos 44, al menos 45 sondas para hibridación con ácidos nucleicos que comprenden un polimorfismo de nucleótido único seleccionado de la Tabla 1, o un polimorfismo de nucleótido único en desequilibrio de ligamiento con uno o más de ellos.

[0174] Los cebadores y sondas para otros SNP pueden incluirse con los kits ejemplificados anteriormente. Por ejemplo, pueden incluirse cebadores y/o sondas para el SNP del cromosoma X (rs5934683) u otros SNP diversos.

EJEMPLOS

EJEMPLO 1 - SNP Indicativos del Riesgo de Cáncer Colorrectal

[0175] Se identificaron 54 SNP asociados al cáncer colorrectal en poblaciones europeas. De ellos, cuatro SNP dentro de 11q12.2 (rs174537, rs4246215, rs174550 y rs1535) están perfectamente correlacionados y pueden ser representados por un haplotipo común (denominado aquí haplotipo 11q12.2). Dos SNP dentro de 19q13.2 (rs1800469 y rs2241714) están perfectamente correlacionados y pueden estar representados por un haplotipo común (denominado aquí haplotipo 19q13.2). Un SNP se encuentra en el cromosoma X (rs5934683) y no se incluyó en la simulación del riesgo de cáncer colorrectal para hombres y mujeres combinados. Dos SNP dentro de 1q41 (rs6687758 y rs6691170) están en desequilibrio de ligamiento. Por lo tanto, se excluyó el rs6691170. Tres SNP dentro de 8q24.21 (rs10505477, rs6983267 y rs7014346) tienen un primo D de 1,0. Por lo tanto, se excluyeron rs10505477 y rs7014346. Dos SNP dentro de 10q24.2 (rs1035209 y rs11190164) tienen un primo D de 0,9. Por lo tanto, se excluyó el rs1035209.

[0176] En consecuencia, se han identificado 45 SNP en total, estando los SNP restantes en desequilibrio de ligamiento con los mismos o en el cromosoma X. Los SNP indicativos de riesgo de cáncer colorrectal se muestran en la Tabla 4. La frecuencia alélica de cada alelo de riesgo y la razón de posibilidades por alelo de riesgo también se muestran en la Tabla 4.

[0177] La frecuencia media de los alelos de riesgo fue de 0,43 (rango de 0,07 a 0,91). La razón de posibilidades media por alelo de riesgo fue de 1,14 (rango de 1,05 a 1,53). El riesgo relativo familiar (FRR; la razón de posibilidades de cáncer colorrectal asociada a tener un familiar de primer grado con cáncer colorrectal) medio que podría atribuirse a cada SNP fue de 1,0040 (rango de 1,0006 a 1,0281), lo que supone el 0,50% (rango de 0,07% a 3,41%) del log FRR total. El FRR combinado que podría atribuirse a los 45 SNP fue de 1,1980, lo que representa el 22,3% del FRR logarítmico total. El FRR estimado no debido a los SNP fue de 1,88.

Tabla 4. SNP asociados al cáncer colorrectal. La tabla indica la nomenclatura del SNP, el gen o genes más cercanos o dentro de la probable diana reguladora del SNP, el genotipo del alelo de riesgo notificado, la frecuencia del alelo de riesgo notificada en los controles, la asociación notificada con el cáncer colorrectal por alelo de riesgo (razón de posibilidades), el riesgo relativo familiar (FRR) atribuible al SNP y la proporción del log FRR debida al SNP. *Gen más cercano o probable diana reguladora del SNP. Los SNP en desequilibrio de ligamiento se muestran entre corchetes [].

Locus	Gen*	SNP	Alelo de riesgo	OR por alelo de riesgo	Frecuencia de alelo de riesgo	FRR	Proporción de log FRR
1p36.2	WNT4; CDC42	rs72847484	T	1.21	0.91	1.003	0.37%
1q25.3	LAMC1	rs10911251	A	1.05	0.54	1.0006	0.07%
1q41	DUSP10; CICP13	rs6687758, [rs6691170]	G	1.09	0.2	1.0012	0.15%
2q32.3	NABP1; MYO18; SDPR	rs11903757	C	1.08	0.36	1.003	0.37%
3p14.1	LRIG1	rs812481	G	1.09	0.58	1.0018	0.22%
3p22.1	RP11; CTNINB1	rs35360328	A	1.14	0.16	1.0023	0.29%
3q26.2	MYNN; TERC	rs10936599	C	1.08	0.75	1.0011	0.14%
4q26	NDST3	rs3987	C	1.36	0.44	1.0235	2.87%
4q32.2	FSTL5	rs35509282	A	1.53	0.09	1.0149	1.83%
5q31.1	PITX1; H2AFY	rs647161	A	1.11	0.67	1.0024	0.30%
6p21.31	CDKN1A	rs1321311	A	1.1	0.23	1.0016	0.20%
8q23.3	EIF3H	rs16892786	C	1.25	0.07	1.0032	0.40%
8q24.21	CCAT2; MYC	rs6983287 [rs10505477, rs7014346]	G	1.21	0.62	1.0091	1.12%
9q24	TPD52L3; UHRF2	rs719725	A	1.19	0.37	1.0011	0.13%
10p13	CUBN	rs10904849	G	1.14	0.68	1.0037	0.46%
10p14	GATA3	rs10795868	G	1.12	0.67	1.0028	0.35%
10q22.3	ZMIZ1; AS1	rs704017	G	1.08	0.57	1.0008	0.10%
10q24.2	SLC25A28; ENTPD7; COX15; CUTC; ABCC2	rs11190164 [rs1035209]	G	1.09	0.29	1.0015	0.19%
10q25	VTG1A	rs12241008	C	1.13	0.09	1.0012	0.15%
11q12.2	FADS1; FEN1	11qhap*; [rs174537, rs4246215, rs174550, rs1535]	G	1.4	0.57	1.0281	3.41%
11q13.4	POLD3	rs3824999	G	1.08	0.5	1.0015	0.18%

(continuación)

5	11q23.1	COLCA2	rs3802842	C	1.11	0.29	1.0022	0.28%
	12p13.32	CCND2	rs3217810	T	1.2	0.16	1.0045	0.55%
	12p13.32	CCND2	rs3217901	G	1.1	0.41	1.0022	0.27%
10	12p13.32	CCND2	rs10774214	T	1.09	0.38	1.0018	0.22%
	12q13.13	DIP2B; ATF1	rs11169552	C	1.09	0.72	1.0015	0.16%
15	12q13.13	LARP4; DIP2B	rs7136702	T	1.06	0.35	1.0008	0.10%
	12q24.12	SH2B3	rs3184504	C	1.09	0.53	1.0019	0.23%
	12q24.21	TBX3	rs59336	T	1.09	0.48	1.0019	0.23%
20	12q24.22	NOS1	rs73206120	G	1.15	0.11	1.0021	0.26%
	14q22.2	BMP4	rs1957636	T	1.06	0.4	1.0014	0.16%
	14q22.2	BMP4	rs4444235	C	1.11	0.46	1.0027	0.33%
25	15q13.3	SCG5; GREM1	rs11632715	A	1.12	0.47	1.0032	0.39%
	15q13.3	SCG5; GREM1	rs16968681	T	1.18	0.09	1.0022	0.28%
	16q22.1	CDH1	rs9629218	G	1.1	0.71	1.0019	0.23%
30	16q24.1	FOXJ1	rs16941835	C	1.15	0.21	1.0032	0.40%
	17q21	STAT3	rs744166	G	1.27	0.55	1.0142	1.74%
	18q21.1	SMAD7	rs4939627	T	1.18	0.52	1.0069	0.84%
35	19q13.11	RHPN2	rs10411210	C	1.15	0.9	1.0018	0.22%
	19q13.2	TMEM91; TGFB1	19qhap*, [rs1800469, rs2241714]	G	1.16	0.49	1.0055	0.68%
40	20p12.3	FERMT1; BMP2	rs2423279	C	1.14	0.3	1.0036	0.44%
	20p12.3	FERMT1; BMP2	rs4813602	G	1.09	0.36	1.0017	0.21%
45	20p12.3	FERMT1; BMP2	rs961253	A	1.12	0.36	1.003	0.36%
	20q13.1	PREX1	rs6066825	A	1.08	0.64	1.0017	0.21%
	20q13.33	LAMA5	rs4925386	C	1.05	0.68	1.0013	0.15%

EJEMPLO 2 - Simulación de Alelos de Riesgo

[0178] Se realizó una simulación para determinar la capacidad del número acumulado de alelos de riesgo de los SNP para discriminar los casos de cáncer colorrectal de los controles y para estimar el riesgo de cáncer colorrectal en función del número de alelos de riesgo utilizando el software PLINK (Purcell et al., 2007) (<http://pngu.mgh.harvard.edu/purcell/plink/>).

[0179] Se simuló una población de 1.000.000 de personas con cáncer colorrectal (casos) y 1.000.000 de personas sin cáncer colorrectal (controles). La distribución de alelos de riesgo de SNP para la población simulada se ajustó a las frecuencias de alelos de riesgo y las razones de posibilidades por alelo notificadas de las asociaciones de cáncer colorrectal. En esta evaluación se asumió un modelo simplista de riesgo en el que la asociación con el cáncer colorrectal para cada SNP era independiente. En este análisis también se asumió que las razones de posibilidades comunicadas para el cáncer colorrectal para cada SNP eran aplicables tanto a hombres como a mujeres y eran constantes con la edad.

[0180] Se evaluó el poder discriminatorio de los SNP para distinguir los casos de los controles utilizando una curva operativa del receptor y estimando el área bajo la curva (la probabilidad de que un caso de cáncer colorrectal seleccionado

aleatoriamente tenga más alelos de riesgo que un control seleccionado aleatoriamente). Se estimaron las razones de posibilidades para el riesgo de cáncer colorrectal para: (i) estar en el quintil más alto y más bajo para el número de alelos de riesgo estar en el quintil medio; (ii) estar en el decil más alto y más bajo para el número de alelos de riesgo frente a estar en la mediana del número de alelos de riesgo; y (iii) por desviación estándar de alelos de riesgo. Los puntos de corte para el número de alelos de riesgo para quintiles y deciles, y la desviación estándar, se basaron en la distribución de alelos de riesgo para los controles.

[0181] Bajo el supuesto de que estas razones de posibilidades eran constantes con la edad e iguales para hombres y mujeres, se estimó el riesgo acumulado de por vida (desde el nacimiento hasta los 70 años) y el riesgo a cinco años para cada categoría de edad de cáncer colorrectal para Australia y EE.UU. por el número de alelos de riesgo SNP. Se asumió que las incidencias específicas por edad de la población australiana y estadounidense eran las incidencias de las personas con la mediana del número de alelos de riesgo. Las incidencias poblacionales de cáncer colorrectal se obtuvieron del Instituto Australiano de Salud y Bienestar, 2015 y de las Estadísticas de Cáncer del Programa de Vigilancia, Epidemiología y Resultados Finales (SEER) (Howlander et al., 1975-2011).

[0182] Se estimó la proporción del riesgo relativo familiar logarítmico (FRR; la razón de posibilidades de cáncer colorrectal asociada a tener un familiar de primer grado con cáncer colorrectal) que podría atribuirse a los alelos de riesgo de los SNP. Se asumió el equilibrio de Hardy-Weinberg para cada SNP, el equilibrio de ligamiento entre los SNP y un modelo multiplicativo para las asociaciones de los SNP con el riesgo de cáncer colorrectal. Más concretamente, dejemos que SNP_1, \dots, SNP_{45} sean los SNP conocidos asociados al cáncer colorrectal y dejemos que factor clínico₁, ..., factor clínico_m sean los desconocidos (nota: estos podrían ser cualquier factor hereditario que contribuya al FRR, pero para simplificar pensamos en ellos como SNP). Entonces, si G_i es una variable aleatoria que da el número de alelos de riesgo en SNP_i para una persona aleatoria de la población, entonces G_1, \dots, G_m son todas variables aleatorias independientes (por equilibrio de ligamiento) y el log-razón de posibilidades para una persona aleatoria es $X_1 + \dots + X_m$ (por el modelo multiplicativo asumido), donde $X_i = G_i \log OR_i$ y OR_i es el razón de posibilidades por alelo para SNP_i . Una formula de Antoniou et al. (2003) ha derivado rigurosamente en Win et al. (2014) se convierte entonces en $\log FRR = \frac{1}{2}[\text{Var}(X_1) + \dots + \text{Var}(X_m)]$.

[0183] Esto muestra que el log FRR es la suma de componentes independientes de los SNP conocidos y desconocidos asociados al cáncer colorrectal. La proporción del log FRR debida a los SNP conocidos es $\frac{1}{22}(\text{Var}(X_1) + \dots + \text{Var}(X_{45}))/\log FRR$ mientras que la proporción debida a los SNP desconocidos es uno menos este valor. Se asumió que el FRR de tener al menos un pariente de primer grado con cáncer colorrectal era de 2,25, basándose en un metaanálisis previo de antecedentes familiares de cáncer colorrectal (Johns et al., 2001) y un cálculo elemental (asumiendo el equilibrio de Hardy-Weinberg) muestra que $\text{Var}(X_i) = 2 p_i (1 - p_i) (\log OR_i)^2$, donde p_i es la frecuencia alélica menor de SNP_i . Utilizando esta estadística, se estimó el riesgo a cinco años de cáncer colorrectal según el número de alelos de riesgo, con y sin antecedentes familiares de cáncer colorrectal.

[0184] El número de alelos de riesgo para las personas simuladas con y sin cáncer colorrectal se muestra en la Figura 1 y puede resumirse como sigue:

- Aquellos con cáncer colorrectal: mediana de 42 alelos de riesgo, rango de 21 a 61 alelos de riesgo, media de 41,6 alelos de riesgo, desviación estándar de 4,2 alelos de riesgo;
- Aquellos sin cáncer colorrectal: mediana 40 alelos de riesgo, rango 20 a 59, media 39,7 alelos de riesgo, desviación estándar 4,2 alelos de riesgo; cuartil superior 44 o más alelos de riesgo; cuartil inferior 36 o menos alelos de riesgo; decil superior 46 o más alelos de riesgo; decil inferior 34 o menos alelos de riesgo) (Figura 1).

[0185] Tener 29 alelos de riesgo correspondía a un riesgo de cáncer colorrectal a lo largo de la vida del 1,4% para una persona de Australia y del 1,0% para una persona de Estados Unidos. Los riesgos respectivos para 36 alelos de riesgo fueron del 2,9% y el 2,0%; para 43 alelos de riesgo fueron del 6,1% y el 4,3%; y para 50 alelos de riesgo fueron del 12,5% y el 8,8% (Figura 1). En comparación con las personas en el quintil medio del número de alelos de riesgo, la razón de posibilidades de cáncer colorrectal fue de 1,81 para las personas en el quintil más alto del número de alelos de riesgo, y de 0,51 para las personas en el quintil más bajo; esto equivale a un riesgo interquintil de 3,55 veces (quintil más alto frente a quintil más bajo). En comparación con las personas con la mediana de 40 alelos de riesgo, la razón de posibilidades de cáncer colorrectal fue de 2,27 para las personas en el decil más alto del número de alelos de riesgo, y de 0,45 para las personas en el decil más bajo; esto equivale a un riesgo 5,04 veces mayor entre deciles (decil más alto frente a decil más bajo). La razón de posibilidades por desviación estándar de los alelos de riesgo fue de 1,57. La curva característica operativa del receptor tenía un área bajo la curva de 0,63.

[0186] Según las tasas de incidencia poblacional de cáncer colorrectal de 2011 en Australia, el riesgo acumulativo medio de cáncer colorrectal hasta los 70 años era del 3,3 %. Para las personas en el quintil más alto por número de alelos de riesgo, el riesgo acumulado fue del 5,9% (11,5% si también tenían un familiar de primer grado con cáncer colorrectal, y 5,5% si no lo tenían) en comparación con el 1,7% para las personas en el quintil más bajo por número de alelos de riesgo (3,2% si también tenían un familiar de primer grado con cáncer colorrectal, y 1,6% si no lo tenían).

[0187] Para las personas en el decil más alto para el número de alelos de riesgo, el riesgo acumulado fue del 7,4% (13,4% si también tenían un pariente de primer grado con cáncer colorrectal, y 6,9% si no lo tenían) en comparación con el 1,5%

para las personas en el decil más bajo para el número de alelos de riesgo (2,8% si también tenían un pariente de primer grado con cáncer colorrectal, y 1,4% si no lo tenían; Figura 2 A, B). Las estimaciones para los hombres fueron por término medio aproximadamente un 13% más altas y para las mujeres fueron por término medio un 16% más bajas que para hombres y mujeres juntos (Figuras 4 y 5).

[0188] El riesgo quinquenal de cáncer colorrectal para la persona media (no afectada previamente) en Australia alcanza el 1% a los 63 años. El mismo riesgo del 1% a 5 años se alcanza aproximadamente 7 años antes para las personas en el quintil más alto por número de alelos de riesgo (y aproximadamente 14 años antes si también tenían antecedentes familiares de cáncer colorrectal), y aproximadamente 10 años antes para las personas en el decil más alto por número de alelos de riesgo (16 años antes si también tenían antecedentes familiares; Figura 2 Paneles C, D y Tabla 5). Por término medio, los varones alcanzaron el umbral de riesgo del 1% 1-2 años antes, y las mujeres alcanzaron el umbral 3-4 años después por término medio que los varones y las mujeres juntos (Tabla 5).

Tabla 5. Edad (años) a la que el riesgo de cáncer colorrectal a 5 años alcanza o supera umbrales del 1%, para diversas categorías de antecedentes familiares de cáncer colorrectal (al menos un familiar de primer grado) y alelos de riesgo de 45 SNP.

Categoría de riesgo	EE.UU.			Australia		
	Todos	Hombres	Mujeres	Todos	Hombres	Mujeres
Población general	70	67	73	63	61	71
Antecedentes familiares (familiar de 1er grado)	58	55	61	53	52	59
Quintil más alto de alelos de riesgo	61	57	62	56	55	62
Decil más alto de alelos de riesgo	58	53	59	53	52	59
Antecedentes familiares y quintil más alto	50	48	52	49	48	55
Antecedentes familiares y decil más alto	48	46	48	47	46	53
Antecedentes familiares y quintil más bajo	71	66	73	63	61	72
Antecedentes familiares y decil más bajo	74	73	80	65	63	76

[0189] Dado que las tasas de incidencia de cáncer colorrectal en la población de los EE.UU. son más bajas (especialmente después de los 50 años en comparación con Australia), los riesgos asociados basados en el número de alelos de riesgo y los antecedentes familiares también son más bajos que los de Australia (Figura 3 Paneles A, B, Figuras 6 y 7). En comparación, el mismo riesgo del 1% se alcanza aproximadamente 9 años antes para las personas en el quintil más alto por número de alelos de riesgo (20 años antes si también tenían antecedentes familiares de cáncer colorrectal), y aproximadamente 12 años antes para las personas en el decil más alto por número de alelos de riesgo (22 años antes si también tenían antecedentes familiares; Figura 3 Paneles C, D y Tabla 5). Por término medio, los varones alcanzaron el umbral de riesgo del 1% entre 3 y 5 años antes, y las mujeres entre 1 y 3 años después que los hombres y las mujeres juntos (Tabla 5).

EJEMPLO 3 - Clasificación de los Sujetos Según el Riesgo de Cáncer Colorrectal

[0190] Se utilizaron simulaciones para cuantificar la utilidad de un panel de 45 SNP asociados al riesgo para categorizar a las personas en función de su riesgo de cáncer colorrectal. Las personas situadas en los extremos del espectro de alelos de riesgo tenían considerablemente más probabilidades de desarrollar cáncer colorrectal (extremo superior) o menos probabilidades (extremo inferior). Dado que la variación total del riesgo asociada a estos SNP en toda la población puede explicar aproximadamente una cuarta parte del FRR total, la fuerza predictiva del perfil SNP aumenta si también se tienen en cuenta los antecedentes familiares de cáncer colorrectal. Dado que la fuerza de la asociación con el cáncer colorrectal para aquellos que se encuentran en el 20% más bajo de la población (para el número de alelos de riesgo de estos SNP) es aproximadamente la inversa del aumento del riesgo asociado con el resto de FRR, las personas que tienen antecedentes familiares de cáncer colorrectal pero que también se encuentran en el quintil más bajo de la población para el número de alelos de riesgo de estos SNP, están en riesgo poblacional.

[0191] Por lo tanto, la medición de estos SNP es un método útil para la evaluación del riesgo de cáncer colorrectal, y puede ser utilizado como una herramienta para determinar quién debe ser recomendado para el cribado del cáncer colorrectal, y en qué intensidad. Por ejemplo, una persona en el 20% superior de la población para alelos de riesgo (al menos 44 alelos) alcanza el riesgo medio de la población a 5 años 9 años antes que la persona media. Por lo tanto, si la persona media alcanza el umbral de riesgo para el cribado de sangre oculta en heces (que recomiendan la mayoría de los programas nacionales de cribado) a los 50 años, entonces una persona con al menos 44 alelos de riesgo alcanza el mismo umbral de riesgo a los 41 años. Las edades para iniciar el cribado mediante colonoscopia en personas con un familiar de primer grado con cáncer colorrectal serían 49 y 47 años para el quintil más alto y el decil más alto de alelos de riesgo, respectivamente. En EE.UU., donde el riesgo poblacional de cáncer colorrectal es inferior al de Australia, el umbral

del 2% para pertenecer al quintil o decil superior y tener antecedentes familiares de cáncer colorrectal se alcanza a los 62 y 59 años respectivamente.

EJEMPLO 4 - Predicción del Riesgo de Cáncer Colorrectal sin Síndrome de Lynch Basada en 45 SNP Independientes Asociados al Riesgo y en Antecedentes Familiares Multigeneracionales

[0192] Se determinó una puntuación de riesgo basada en los antecedentes familiares que proporciona un riesgo de cáncer colorrectal a 5 años ajustado por edad y transformado logarítmicamente basado en datos de cáncer colorrectal multigeneracional utilizando un modelo mixto de genes principales - poligénico (CRiPT). Esta evaluación del riesgo clínico se combinó con la puntuación de riesgo basada en los 45 SNP enumerados en la Tabla 4. Los inventores utilizaron la regresión logística para estimar la razón de posibilidades por desviación estándar ajustada (OPERA) (Dite et al., 2016) para cada puntuación con riesgo de cáncer colorrectal.

[0193] La puntuación basada en el SNP, la puntuación basada en los antecedentes familiares y las puntuaciones combinadas basadas en el SNP y los antecedentes familiares se asociaron con el riesgo de cáncer colorrectal con OPERAs de 1,40 (intervalo de confianza [IC] del 95%, 1,24-1,58), 1,39 (1,26-1,53) y 1,59 (1,42-1,79), respectivamente. Equivalen a cocientes de riesgo intercuartílicos (riesgo en el 25% más alto de la población para la puntuación de riesgo dividido por el riesgo en el 25% más bajo de la población) de 2,4, 2,3 y 3,2. La puntuación de riesgo combinada dio mejores resultados que las puntuaciones basadas en el SNP y en los antecedentes familiares (ambas $P < 0,001$). En el caso de las personas con antecedentes familiares moderadamente fuertes que las sitúan en un riesgo aproximadamente 4 veces mayor (similar al de tener dos familiares de primer grado diagnosticados de cáncer colorrectal mayores de 50 años), estas estimaciones predicen que los que se encuentran en el cuartil superior (25%) para las puntuaciones SNP tienen un riesgo más de 6 veces superior al de la población, mientras que los que se encuentran en el cuartil inferior tienen un riesgo inferior a 2,5 veces el de la población.

[0194] Así, la combinación de información sobre SNP con la historia familiar multigeneracional mejoró la capacidad de predicción del cáncer colorrectal en aproximadamente un 40%. Por lo tanto, dado que podría reclasificar el tratamiento clínico de aproximadamente la mitad de estas personas, esta nueva medida de riesgo combinada puede utilizarse para orientar mejor el cribado del cáncer colorrectal en función del riesgo.

[0195] La presente solicitud reivindica prioridad de AU 2016900254 presentada el 28 de enero de 2016 y 2016903246 presentada el 16 de agosto de 2016,

[0196] Cualquier referencia a documentos, actos, materiales, dispositivos, artículos o similares que se haya incluido en la presente especificación tiene como único propósito proporcionar un contexto para la presente divulgación. No debe considerarse como una admisión de que alguno o todos estos asuntos formen parte del estado de la técnica o fueran de conocimiento general en el campo relevante para la presente divulgación tal y como existía antes de la fecha de prioridad de cada reivindicación de esta solicitud.

REFERENCIAS

[0197]

- Ait Ouakrim et al. (2012) Cancer Prev Res. (Phila) 5:240-247.
 Antoniou et al. (2003) Genet Epidemiol. 25:190-202.
 Ausubel et al. (editores) (1998), Current Protocols in Molecular Biology, Greene Pub. Associates y Wiley-Interscience (incluidas todas las actualizaciones hasta la fecha).
 Brenner et al. (2014) BMJ 348 g2467.
 Brown (editor) (1991), Essential Molecular Biology: A Practical Approach.,
 Coligan et al. (editores) Current Protocols in Immunology, John Wiley & Sons (incluidas todas las actualizaciones hasta la fecha).
 Devlin and Risch (1995) Genomics. 29: 311-322.
 Dite et al. (2016) Cancer Epidemiol Biomarkers Prev 25:359-365.
 Glover y Hames (editores) (1995 y 1996) DNA Cloning: A Practical Approach, Volúmenes 1-4, IRL Press.
 Harlow y Lane (editores) (1988) Antibodies: A Laboratory Manual, Cold Spring Harbour Laboratory.
 Hartl (1981) A Primer of Population Genetics Washington University, Saint Louis Sinauer Associates, Inc. Sunderland, Mass. ISBN: 0-087893-271-2.
 Hewitson et al. (2007) The Cochrane database of systematic reviews doi:10.1002/14651858.CD001216.pub2(1), CD001216.
 Johns et al (2001) Gastroenterol. 96:2992-3003.
 Lynch y Walsh (1998) Genetics and Analysis of Quantitative Traits, Sinauer Associates, Inc. Sunderland Mass. ISBN 0-87893-481-2.
 Mavaddat et al. (2015) J Natl Cancer Inst 107:djv036.
 Pencina et al. Statistics in Medicine 2008; 27(10): 157-172.
 Perbal (2000) A Practical Guide to Molecular Cloning, John Wiley and Sons (1984)
 Purcell et al. (2007) Am J Hum Genet. 81:559-575.
 Purcell et al. (2007) (<http://pngu.mgh.harvard.edu/purcell/plink/>).

Sambrook et al. (1989) Clonación molecular: A Laboratory Manual, Cold Spring Harbour Laboratory Press.

Slatkin y Excoffier (1996) Heredity 76: 377-383.

España et al. (2012) Hum Mol Genet. 21:934-946.

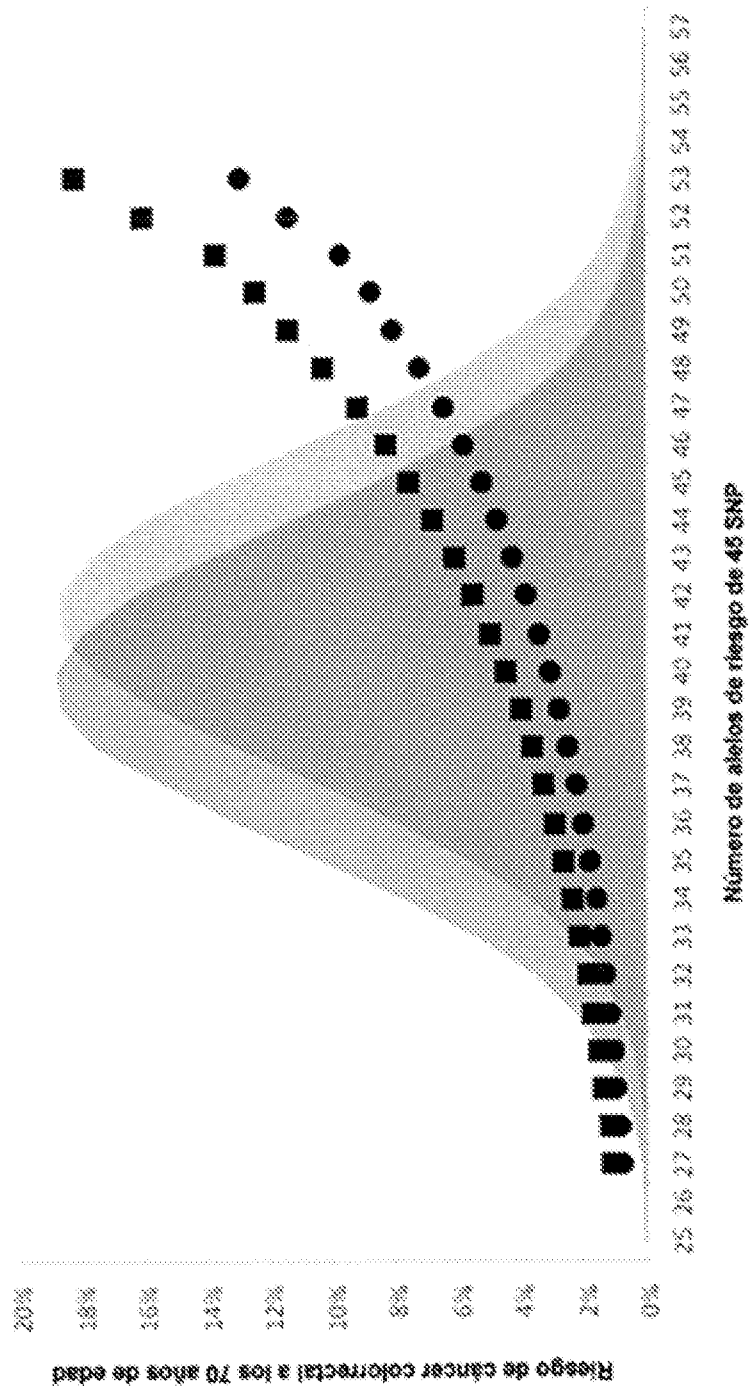
Tijssen (1993) Laboratory Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Acid Probes Elsevier, Nueva York.

Usher-Smith et al. (2015) Cancer Prev Res 9: 13-26.

Win et al. (2014) Gastroenterology 146:1208-1211, e1201-1205.

REIVINDICACIONES

- 5 1. Un método para evaluar el riesgo de un sujeto humano de desarrollar cáncer colorrectal que comprende realizar una evaluación del riesgo genético del sujeto, donde la evaluación del riesgo genético implica detectar, en una muestra biológica derivada del sujeto, la presencia de al menos los 45 polimorfismos de nucleótido único proporcionados en la Tabla 1.
- 10 2. El método de la reivindicación 1, que comprende además realizar una evaluación del riesgo clínico del sujeto y combinar la evaluación del riesgo genético con la evaluación del riesgo clínico para obtener el riesgo de un sujeto humano de desarrollar cáncer colorrectal.
- 15 3. El método de la reivindicación 1 o de la reivindicación 2, en el que el sujeto ha tenido una prueba de sangre oculta en heces positiva.
4. El método de cualquiera de las reivindicaciones 1 a 3, en el que el sujeto tiene antecedentes familiares de cáncer colorrectal y tiene al menos 30 años.
- 20 5. El método de cualquiera de las reivindicaciones 1 a 4, en el que el sujeto es varón.
6. El método de cualquiera de las reivindicaciones 1 a 5, en el que los resultados de la evaluación del riesgo indican que el sujeto debe inscribirse en un programa de cribado o someterse a un cribado más frecuente.
- 25 7. El método de cualquiera de las reivindicaciones 1 a 6, en el que el rendimiento del método se **caracteriza por** un área bajo la curva (AUC) de al menos aproximadamente 0,63.
8. Un método para determinar la necesidad de pruebas diagnósticas rutinarias de un sujeto humano para cáncer colorrectal que comprende evaluar el riesgo del sujeto de desarrollar cáncer colorrectal utilizando el método de cualquiera de las reivindicaciones 1 a 7.
- 30 9. Un método de cribado de cáncer colorrectal en un sujeto humano, el método comprende evaluar el riesgo del sujeto de desarrollar cáncer colorrectal utilizando el método de cualquiera de las reivindicaciones 1 a 7, y cribado rutinario de cáncer colorrectal en el sujeto si se evalúa que tiene riesgo de desarrollar cáncer colorrectal.
- 35 10. Un método implementado por ordenador para evaluar el riesgo de un sujeto humano de desarrollar cáncer colorrectal, el método operable en un sistema informático que comprende un procesador y una memoria, el método comprende:
40 recibir datos de riesgo genético del sujeto, en los que los datos de riesgo genético se obtuvieron detectando, en una muestra biológica derivada del sujeto, la presencia de al menos los 45 polimorfismos de un solo nucleótido proporcionados en la Tabla 1;
procesar los datos para determinar el riesgo del sujeto humano de desarrollar cáncer colorrectal;
emitir el riesgo del sujeto humano de desarrollar cáncer colorrectal.
- 45 11. El método implementado por ordenador de la reivindicación 10, que comprende además recibir datos de riesgo clínico del sujeto;
50 procesar los datos para combinar los datos de riesgo clínico con los datos de riesgo genético para obtener el riesgo del sujeto de desarrollar cáncer colorrectal;
emitir el riesgo del sujeto de desarrollar cáncer colorrectal.
12. El método implementado por ordenador de la reivindicación 10 o de la reivindicación 11, en el que los datos de riesgo del sujeto se reciben desde una interfaz de usuario acoplada al sistema informático.
- 55



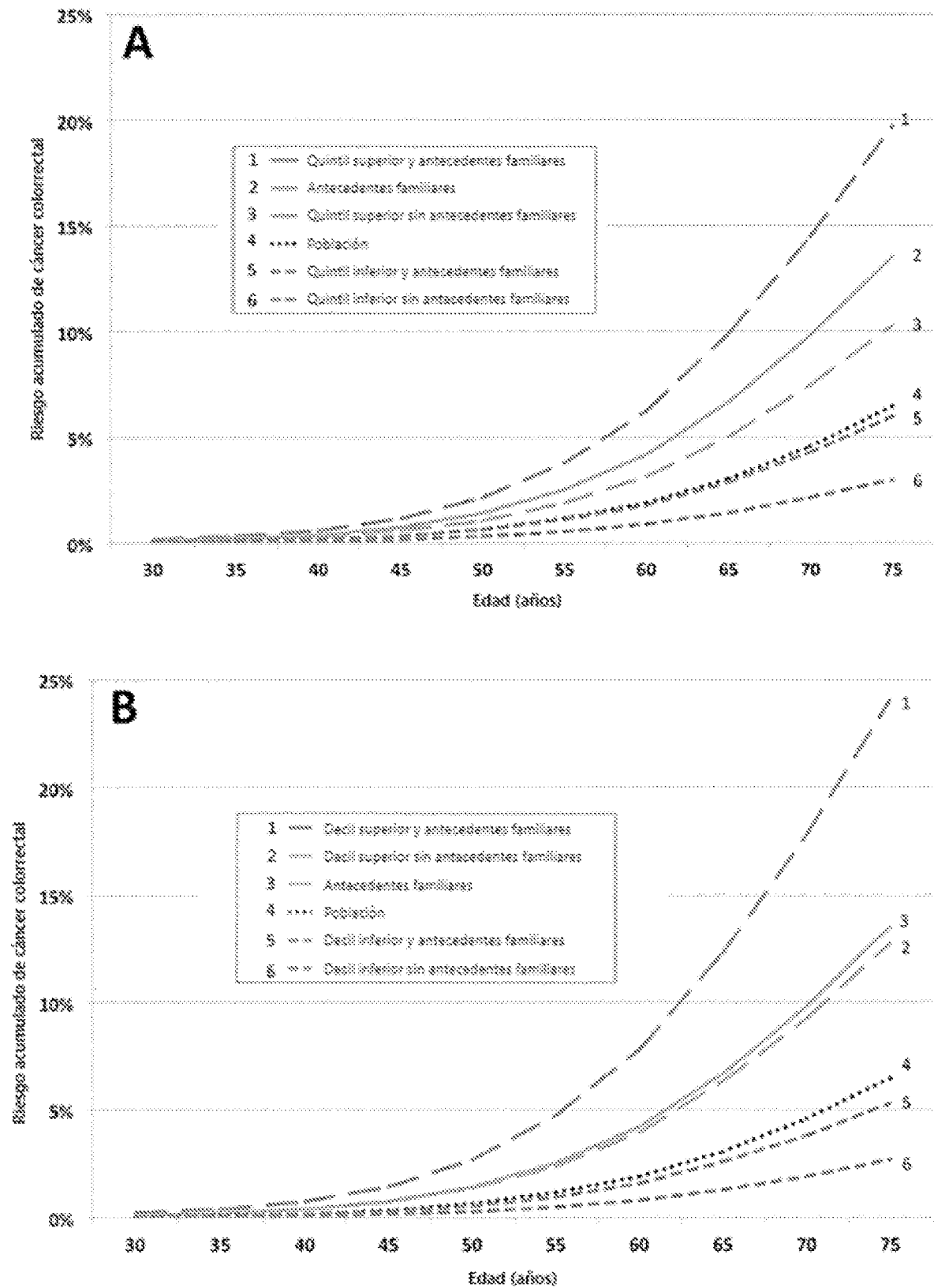


FIGURA 2-1

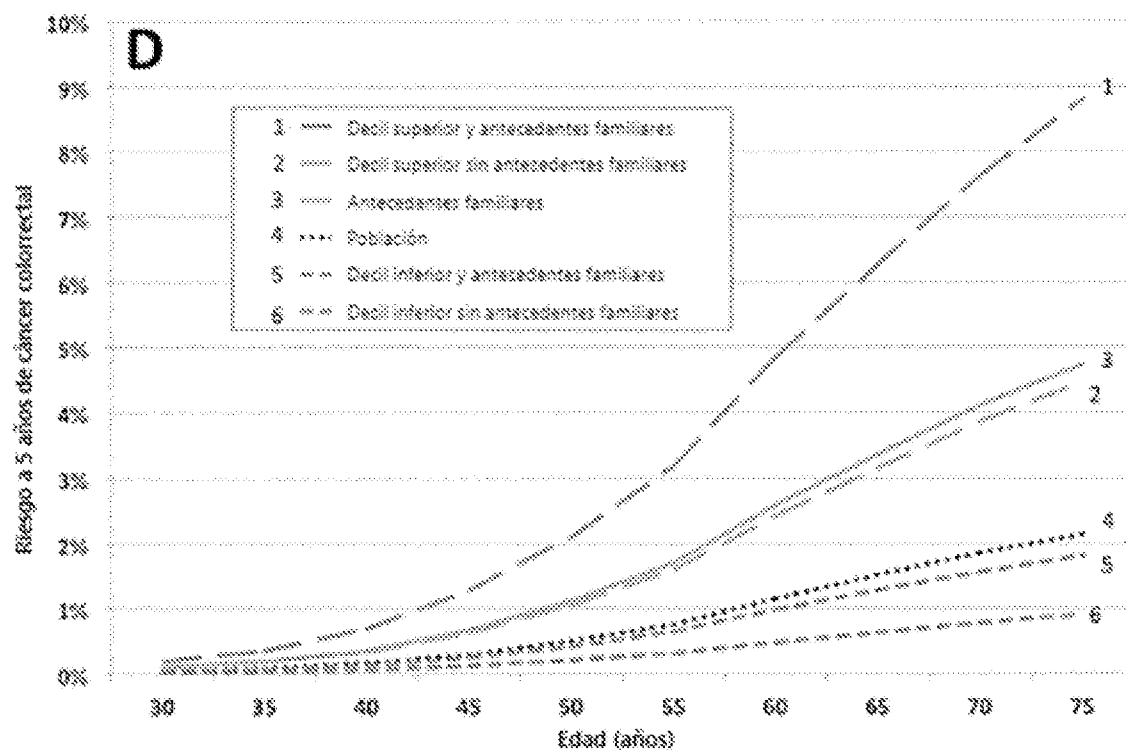
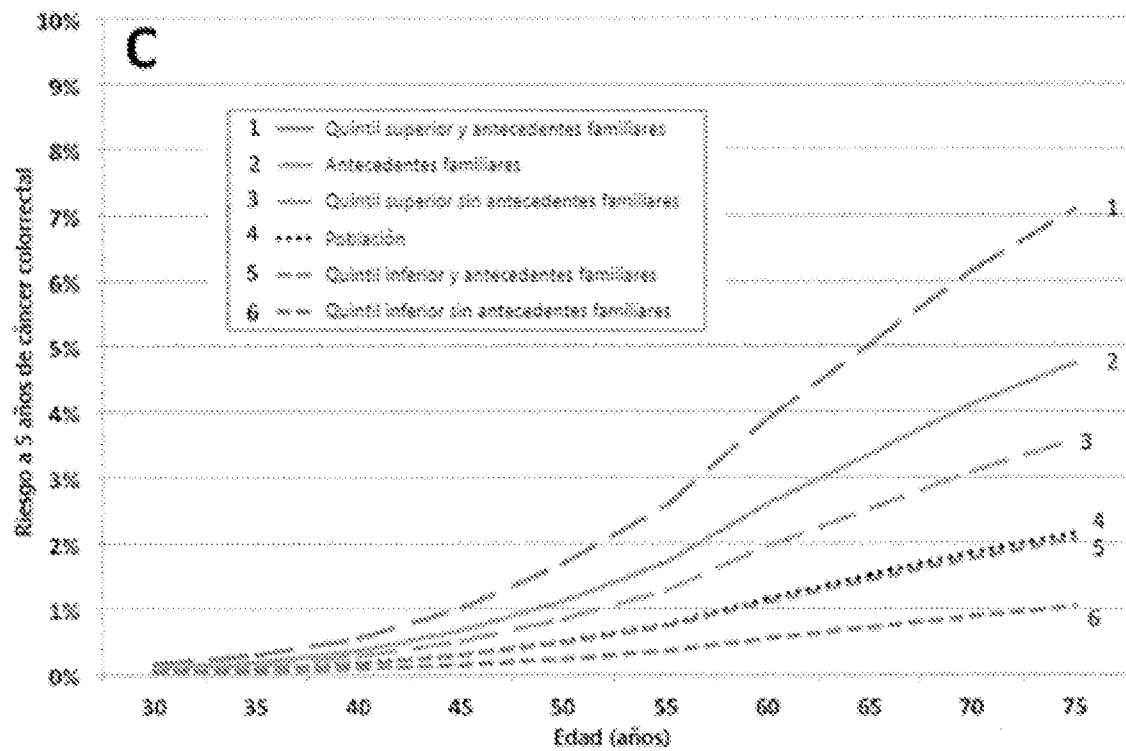


FIGURA 2-2

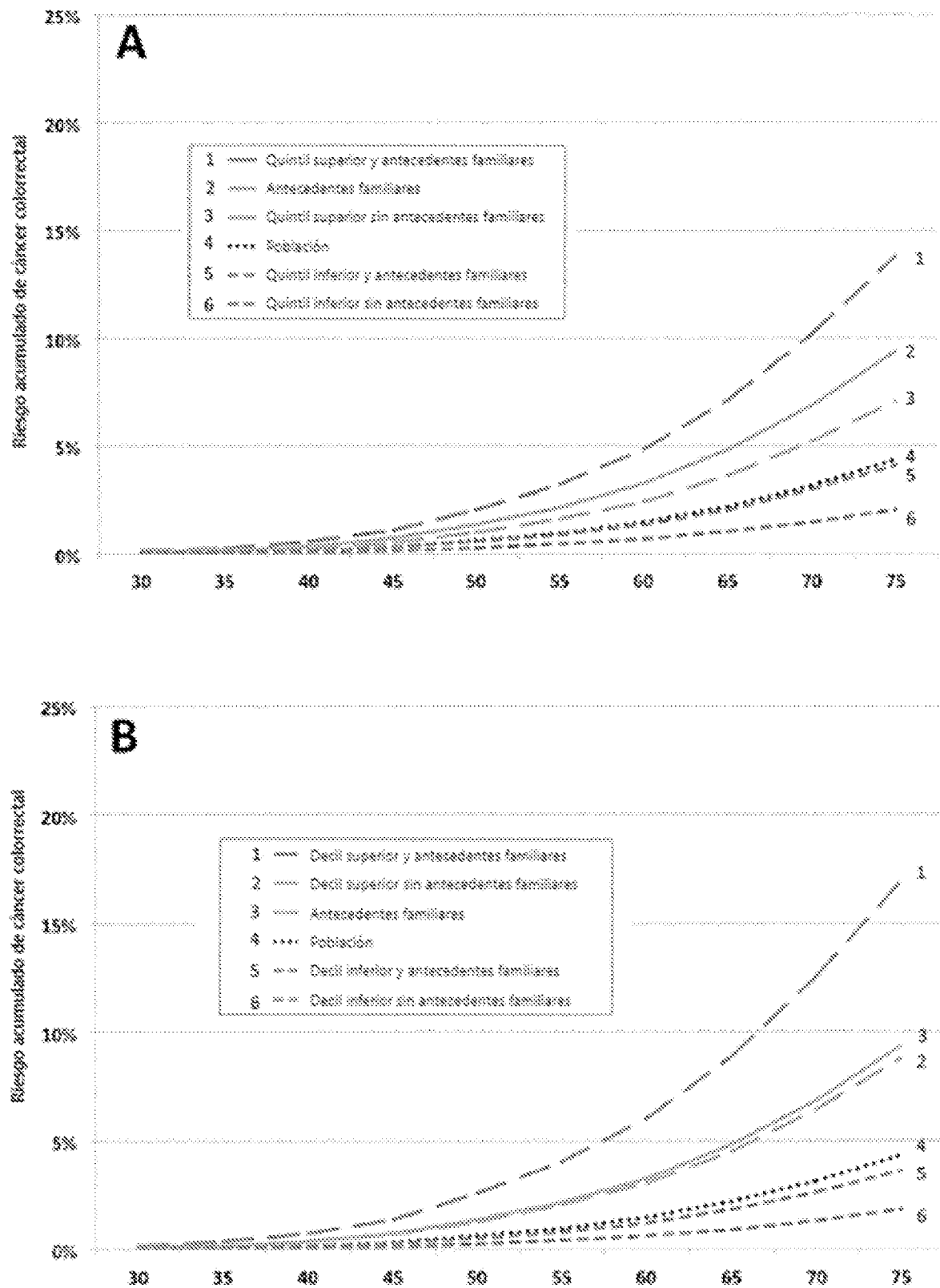


FIGURA 3-1

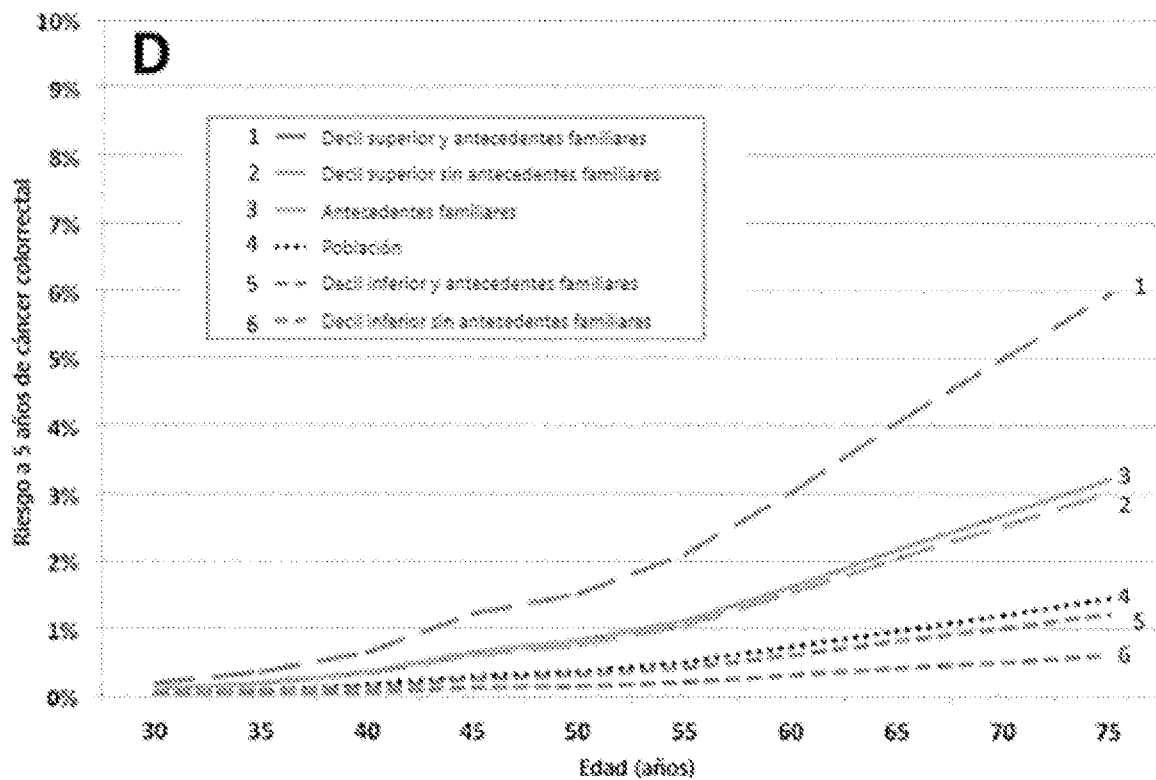
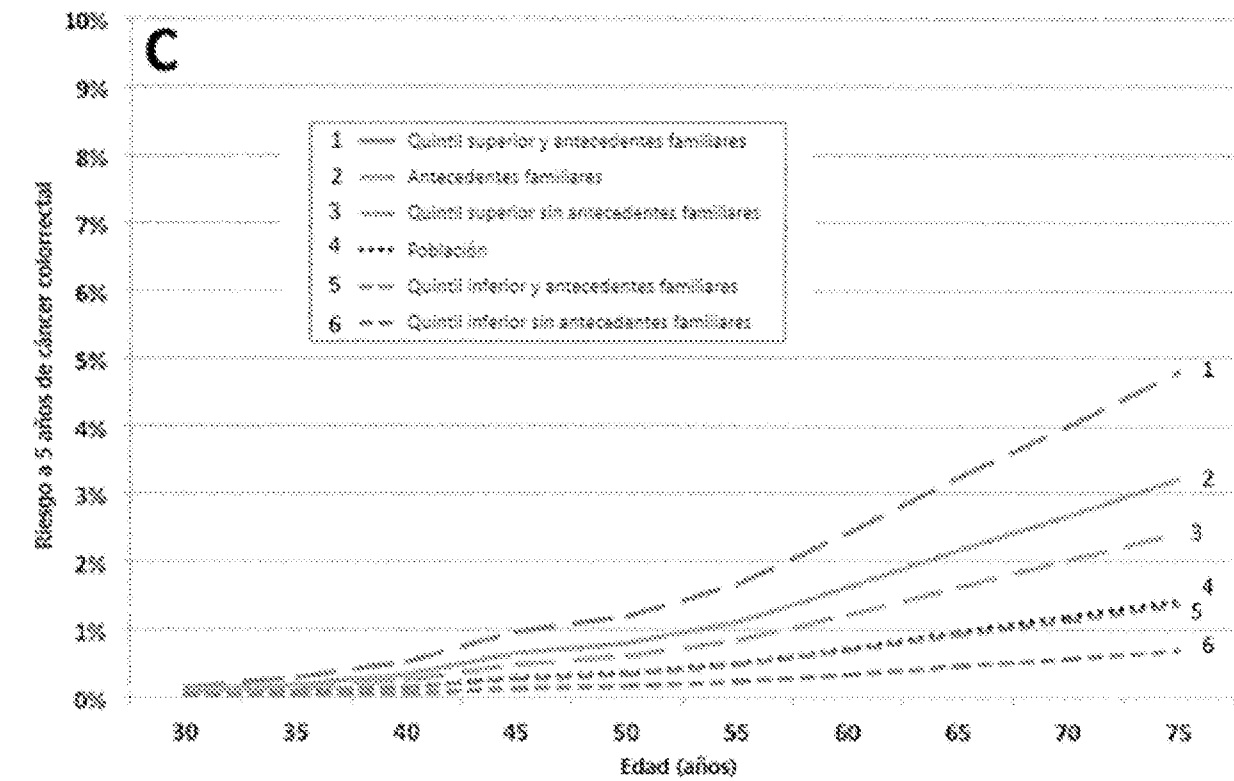


FIGURA 3-2

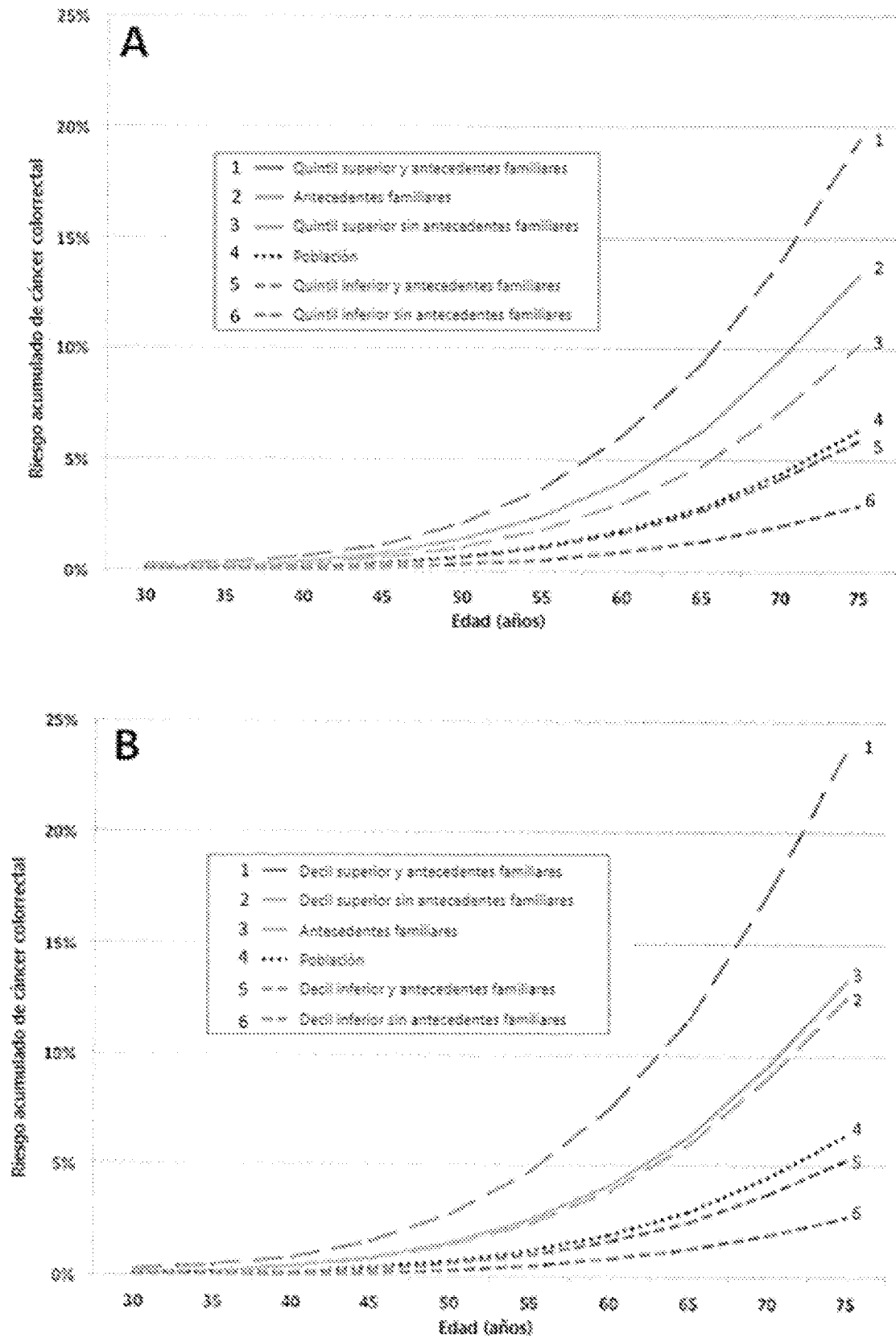


FIGURA 4-1

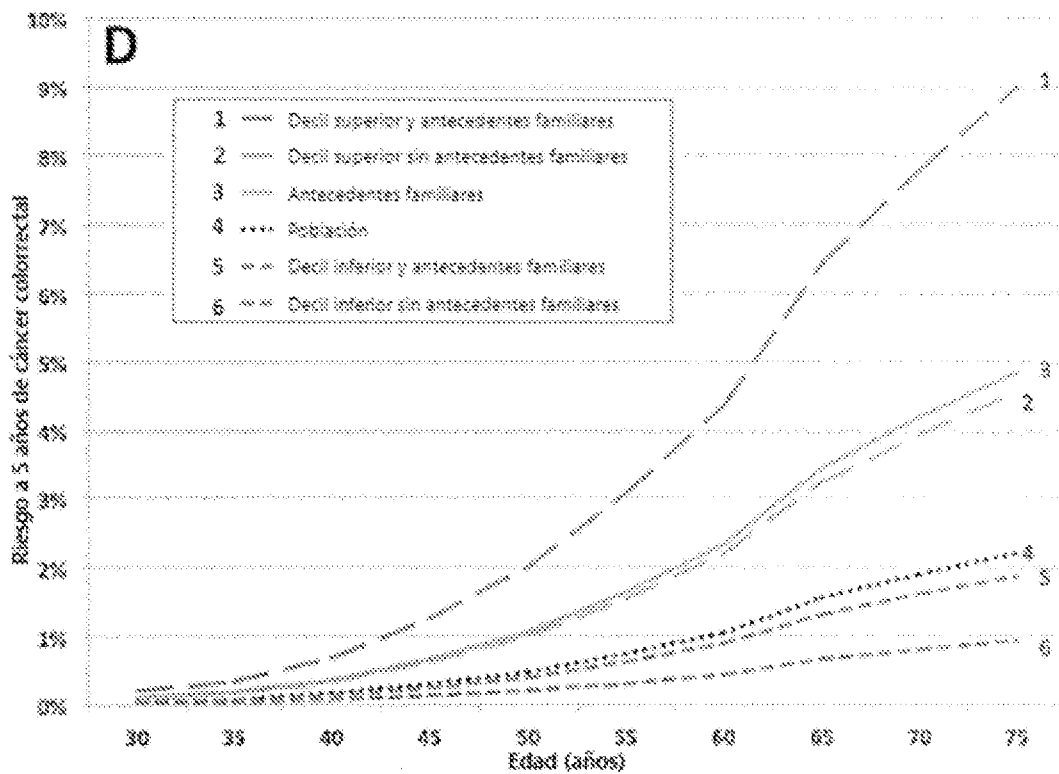
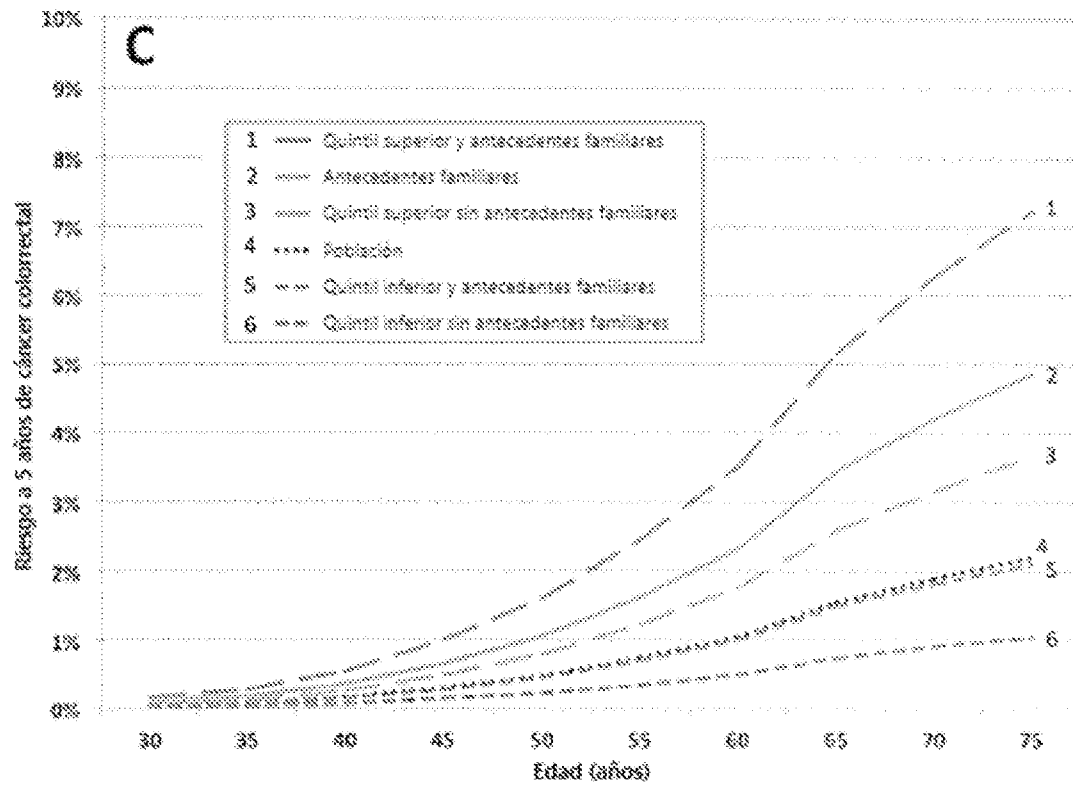


FIGURA 4-2

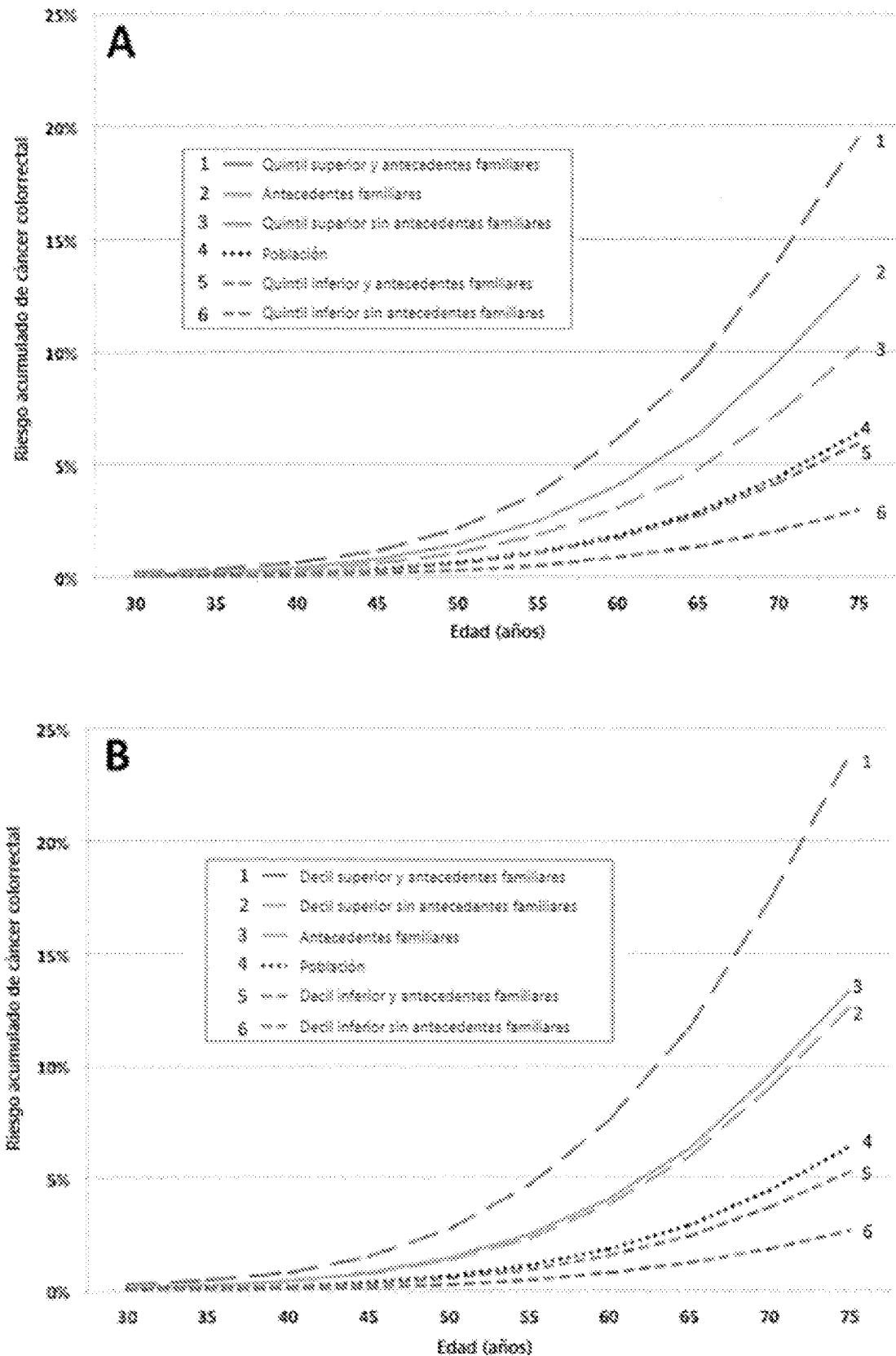


FIGURA 5-1

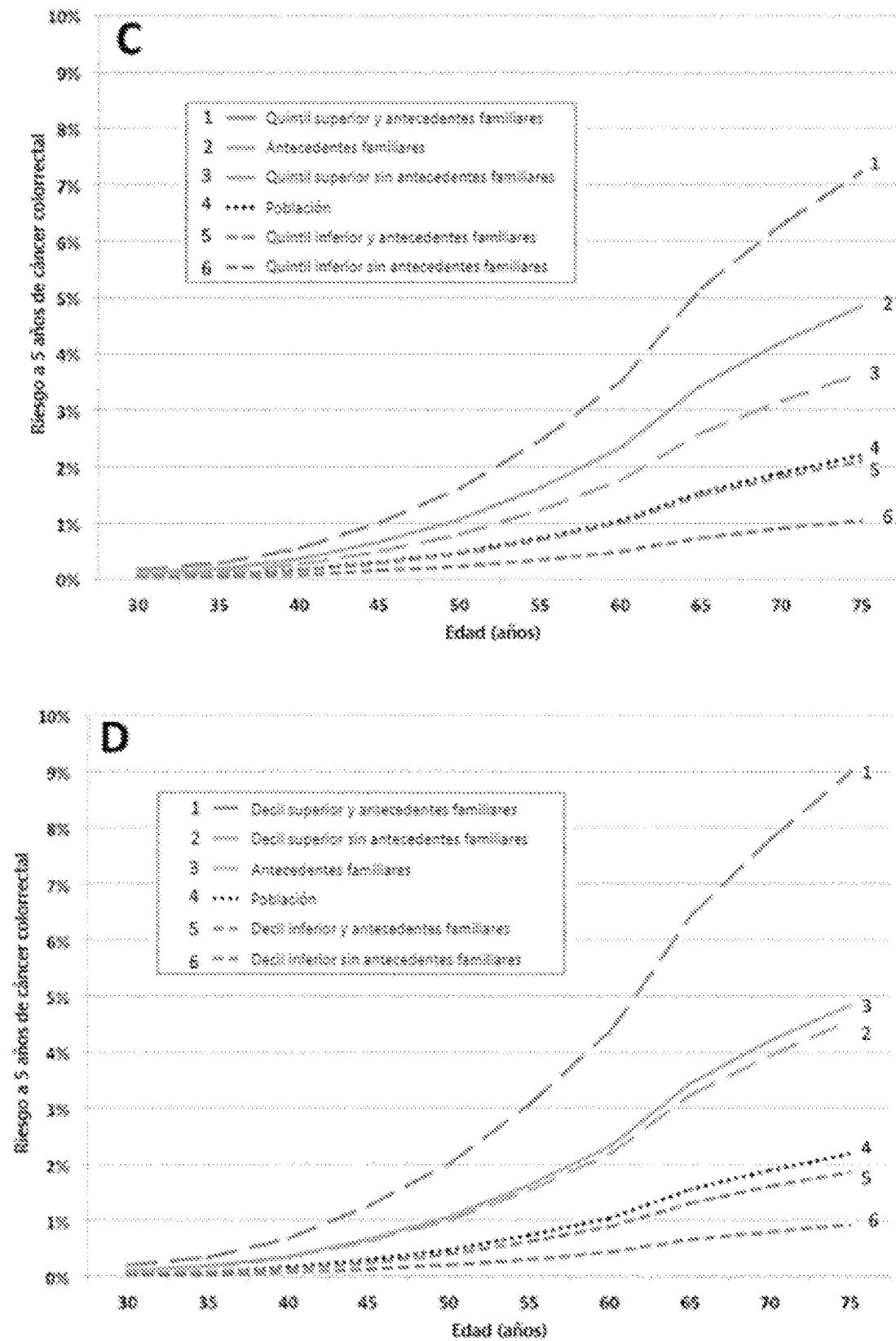


FIGURA 5-2

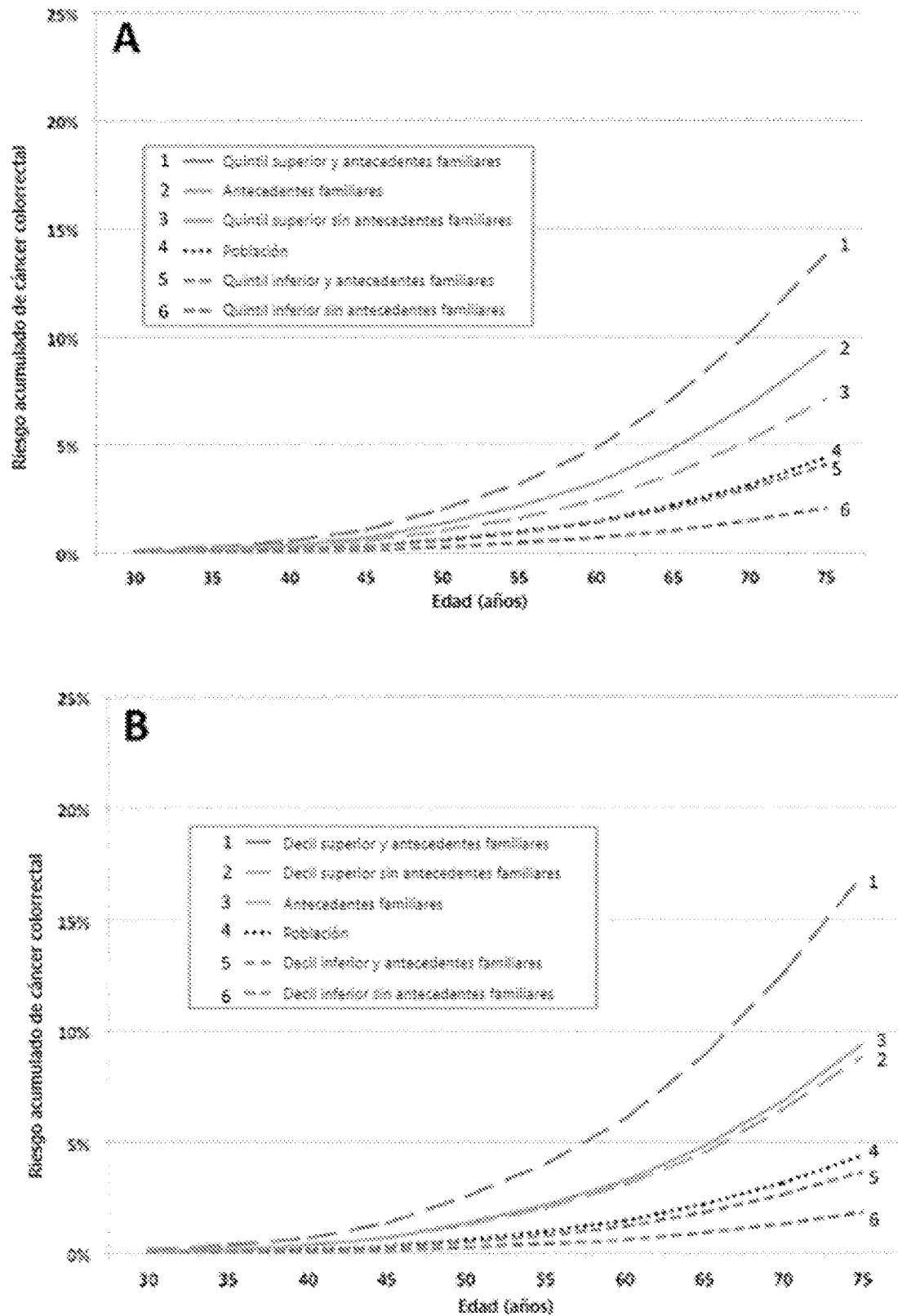


FIGURA 6-1

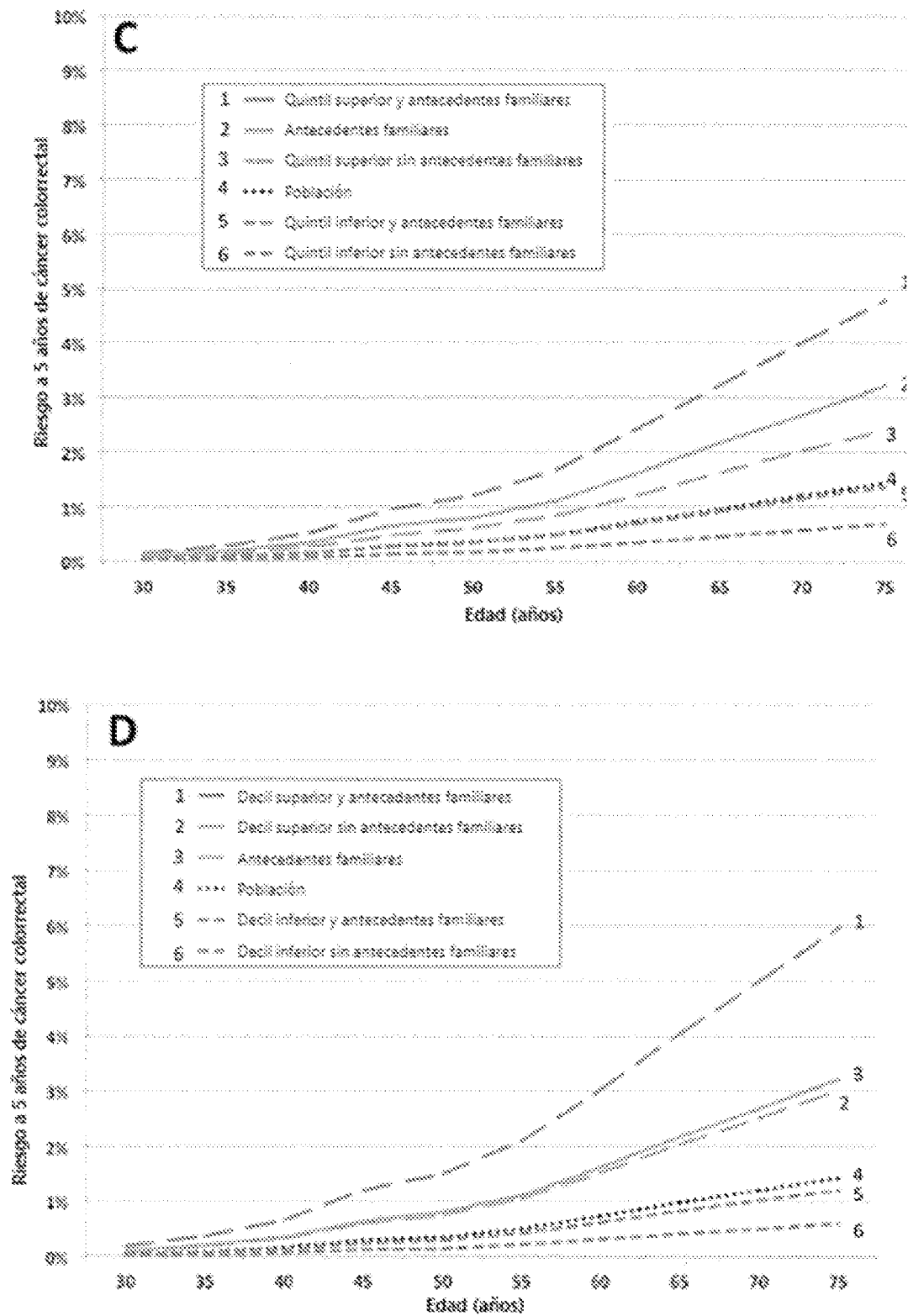


FIGURA 6-2

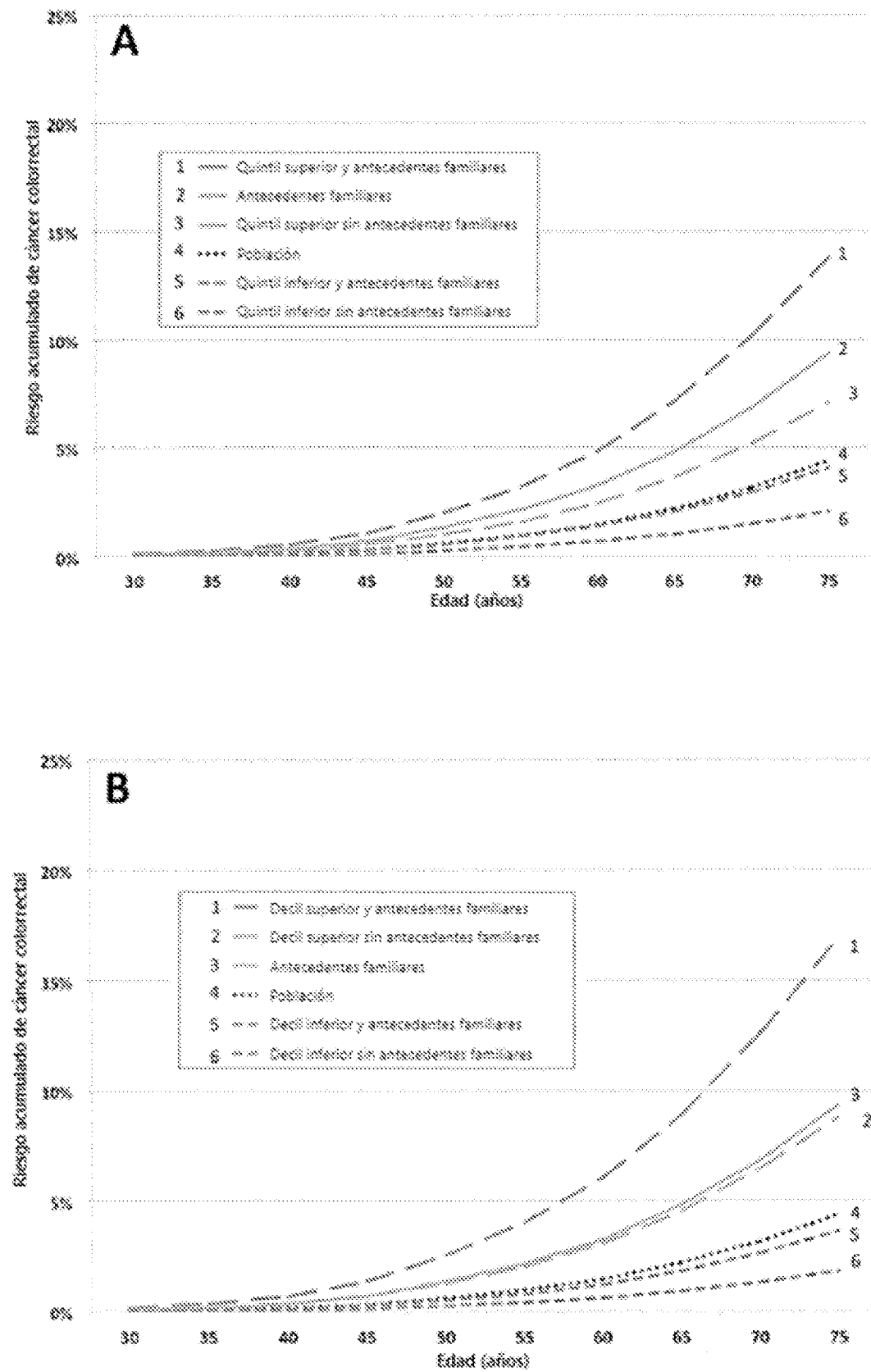


FIGURA 7-1

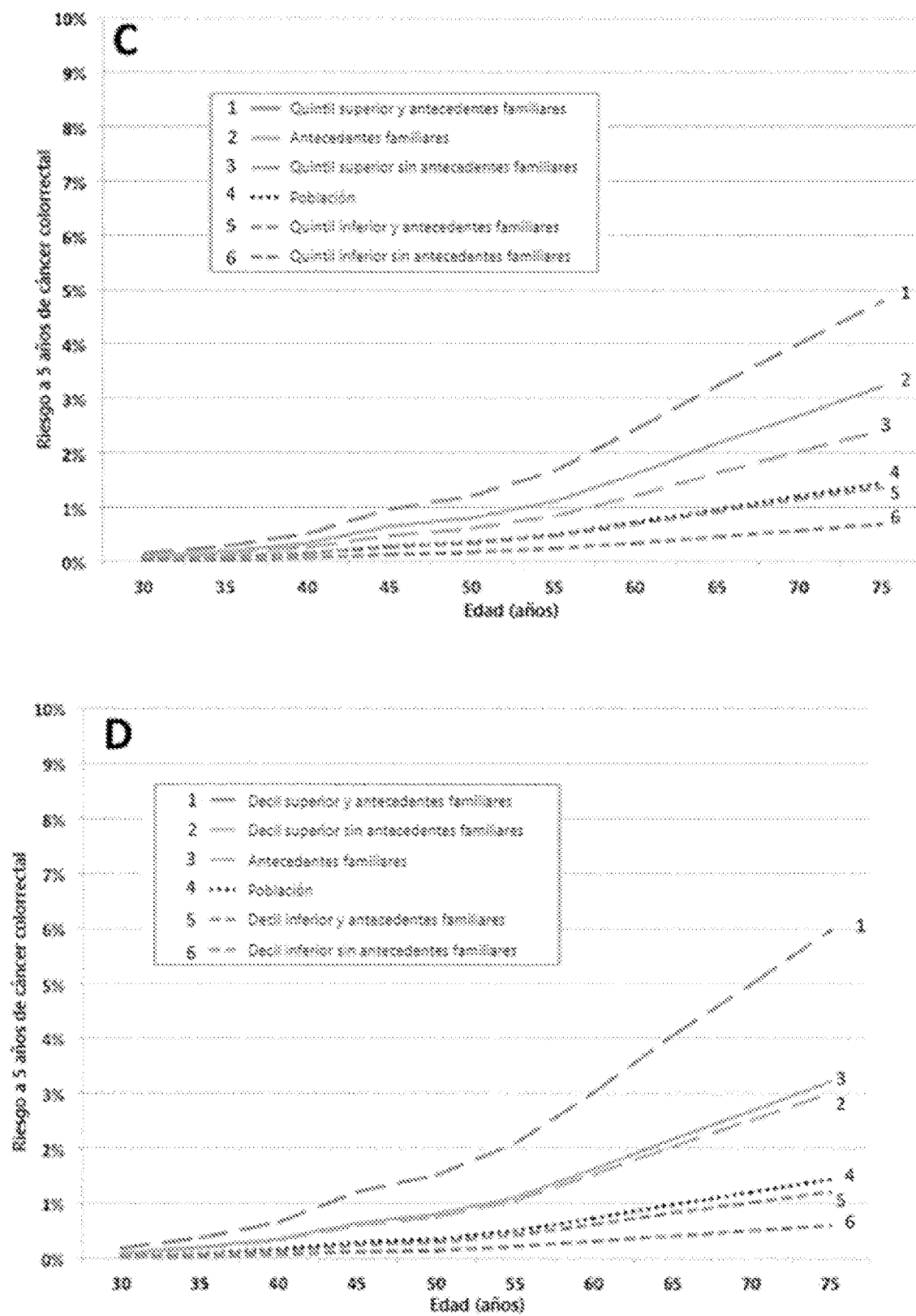


FIGURA 7-2