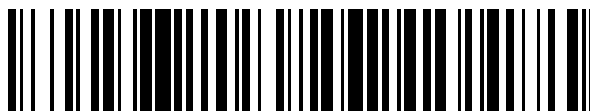


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 573 133**

51 Int. Cl.:

C12Q 1/68

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **16.12.2011** **E 11804757 (0)**

97 Fecha y número de publicación de la concesión europea: **09.03.2016** **EP 2652145**

54 Título: **Conjunto de sondas oligonucleotídicas y métodos para realizar el perfil de microbiota**

30 Prioridad:

16.12.2010 GB 201021399

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

06.06.2016

73 Titular/es:

GENETIC ANALYSIS AS (100.0%)

Pb 4239 Nydalen

0401 Oslo, NO

72 Inventor/es:

VEBØ, HEIDI;

RUDI, KNUT y

SEKELJA, MONIKA

74 Agente/Representante:

ISERN JARA, Jorge

ES 2 573 133 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

Conjunto de sondas oligonucleotídicas y métodos para realizar el perfil de microbiota

La presente divulgación se refiere a un conjunto de sondas oligonucleotídicas y su uso para realizar el perfil de la microbiota del tracto gastrointestinal (GI). Pueden por lo tanto identificarse perfiles de microbiota del tracto GI característicos de una enfermedad o afección o el riesgo de desarrollar una enfermedad o afección. Estos perfiles de microbiota característicos pueden usarse después en el diagnóstico o la supervisión de dichas enfermedades y afecciones. El conjunto de sondas puede proporcionarse en forma de kit.

El tracto GI, también denominado el tracto digestivo o canal alimentario (y cuyos términos pueden usarse indistintamente con tracto GI) es la serie continua de órganos que comienzan en la boca y terminan en el ano. En toda su longitud el tracto GI está colonizado por microorganismos de diversas especies diferentes. En su conjunto el contenido de microorganismos del tracto GI es la microbiota del tracto GI y las cantidades relativas de los microorganismos constituyentes pueden considerarse un perfil de la microbiota. También puede determinarse la microbiota y perfiles de microbiota de diferentes regiones del tracto GI.

Se cree que muchas enfermedades y afecciones, o estadios de las mismas, están ligadas a perfiles característicos de la microbiota del tracto GI, o regiones del mismo. En algunos casos la enfermedad o afección puede estar provocada por, o se exacerba por, alteraciones del perfil de la microbiota del tracto GI. En otros casos la enfermedad o afección provoca, o en algunos mecanismos da como resultado, la presentación de un perfil particular de la microbiota del tracto GI. En consecuencia, analizando los perfiles de microbiota en muestras del tracto GI, puede proporcionarse información que permite el diagnóstico o supervisión de una enfermedad o afección, o que permite una evaluación del riesgo de desarrollar una enfermedad o afección, que se ha determinado que se caracteriza por un perfil de microbiota particular.

Las enfermedades y afecciones que afectan al tracto GI muy probablemente tienen como resultado perfiles de microbiota característicos, por ejemplo enfermedad inflamatoria del intestino (EII), enfermedad de Crohn (EC), colitis ulcerosa (CU), síndrome del intestino irritable (SII) y cánceres del tracto GI (por ejemplo cáncer de la boca, la faringe, el esófago, el estómago, el duodeno, el yeyuno, el íleon, el ciego, el colon, el recto o el ano) y también existen pruebas de relaciones entre la microbiota del tracto GI y enfermedades y afecciones que se considera que no están relacionadas con el tracto GI, por ejemplo las enfermedades atópicas, por ejemplo, eccema, asma, dermatitis atópica, conjuntivitis alérgica, rinitis alérgica y alergias alimentarias; trastornos metabólicos, por ejemplo diabetes mellitus (tipo 1 y tipo 2), obesidad y síndrome metabólico; trastornos neurológicos, por ejemplo depresión, esclerosis múltiple, demencia y enfermedad de Alzheimer y autismo.

Se ha identificado ahora un conjunto de sondas que pueden analizar con un alto grado de sensibilidad y precisión las cantidades relativas de bacterias constituyentes claves de la microbiota del tracto GI y de este modo proporcionan perfiles de la microbiota del tracto GI que son suficientemente detallados y precisos para ser característicos de diversas enfermedades o afecciones o el riesgo de desarrollar diversas enfermedades o afecciones. En consecuencia el conjunto de sondas recién identificado es una herramienta de diagnóstico potente de alta sensibilidad y precisión.

Por lo tanto, en un aspecto se proporciona un conjunto de sondas oligonucleotídicas, comprendiendo dicho conjunto:

(a) un oligonucleótido que comprende una secuencia de nucleótidos seleccionada de ACGCTTGACCCCT (SEQ ID NO 1), la secuencia complementaria de la misma (AGGGTGCAAGCGT; SEQ ID NO 27) o una secuencia capaz de hibridar con una de las secuencias en condiciones de alta rigurosidad;

(b) un oligonucleótido que comprende una secuencia de nucleótidos seleccionada de CGATCCGAAAACCTTCTTCACT (SEQ ID NO 2), la secuencia complementaria de la misma (AGTGAAGAAGGTTTTCGGATCG; SEQ ID NO 28) o una secuencia capaz de hibridar con una de las secuencias en condiciones de alta rigurosidad;

(c) un oligonucleótido que comprende una secuencia de nucleótidos seleccionada de GGACAACGCTTGCCAC (SEQ ID NO 3), la secuencia complementaria de la misma (GTGGCAAGCGTTGTCC; SEQ ID NO 29) o una secuencia capaz de hibridar con una de las secuencias en condiciones de alta rigurosidad;

(d) un oligonucleótido que comprende una secuencia de nucleótidos seleccionada de CGTAGGCGGTTTCGTCGCGT (SEQ ID NO 4), la secuencia complementaria de la misma (ACGCGACGAACCGCCTACG; SEQ ID NO 30) o una secuencia capaz de hibridar con una de las secuencias en condiciones de alta rigurosidad; y

(e) uno o más oligonucleótidos que comprenden una secuencia de nucleótidos seleccionada de las enumeradas en la Tabla 1 o una secuencia capaz de hibridar con cualquier secuencia de nucleótidos enumerada en la Tabla 1 en condiciones de alta rigurosidad; y opcionalmente

(f) uno o más oligonucleótidos que comprenden una secuencia de nucleótidos seleccionada de las enumeradas en la Tabla 2 y la Tabla 3 o una secuencia capaz de hibridar con cualquier secuencia de nucleótidos enumerada en la Tabla 2 y la Tabla 3 en condiciones de alta rigurosidad.

En realizaciones preferidas el componente (f) está presente en el conjunto de sondas y es uno o más oligonucleótidos que comprenden una secuencia de nucleótidos seleccionada de las enumeradas en la Tabla 2.

5 Todas y cada una de las combinaciones de las diversas opciones individuales para cada componente se contemplan específicamente y se desvelan por la presente.

Por lo tanto, en ciertas otras realizaciones el conjunto de sondas oligonucleotídicas comprende componentes (a) a (d) y opcionalmente componente (f) todos como se ha definido anteriormente y al menos uno de

- 10 (i) un oligonucleótido que comprende una secuencia de nucleótidos seleccionada de SEQ ID NO 5, SEQ ID NO 31 o una secuencia capaz de hibridar con una de las secuencias en condiciones de alta rigurosidad,
 (ii) un oligonucleótido que comprende una secuencia de nucleótidos seleccionada de SEQ ID NO 6, SEQ ID NO 32 o una secuencia capaz de hibridar con una de las secuencias en condiciones de alta rigurosidad,
 15 (iii) un oligonucleótido que comprende una secuencia de nucleótidos seleccionada de SEQ ID NO 7, SEQ ID NO 33 o una secuencia capaz de hibridar con una de las secuencias en condiciones de alta rigurosidad,
 (iv) un oligonucleótido que comprende una secuencia de nucleótidos seleccionada de SEQ ID NO 8, SEQ ID NO 34 o una secuencia capaz de hibridar con una de las secuencias en condiciones de alta rigurosidad.

20 El conjunto de sondas puede comprender, y típicamente comprenderá, más de una copia de cada especie de sonda oligonucleotídica seleccionada.

Pueden estar presentes en el conjunto de sondas sondas oligonucleotídicas adicionales. Preferentemente las sondas oligonucleotídicas contribuirán a la información sobre el contenido de microbiota del tracto GI que puede proporcionar el conjunto de sondas. Esto puede ser porque las sondas adicionales proporcionan información positiva 25 sobre la microbiota del tracto GI o proporcionando información que puede actuar como control para una o más de las otras sondas en el conjunto de sondas o información normalizada que podría permitir la cuantificación de la información obtenida de una o más de las otras sondas en el conjunto de sondas. Las sondas adicionales pueden dirigirse a las mismas o diferentes bacterias como una o más de las sondas del conjunto de sondas definido anteriormente.

30 También se describe en el presente documento una sonda oligonucleotídica que comprende una secuencia de nucleótidos seleccionada de una cualquiera de SEQ ID NO 1-52 o una secuencia de nucleótidos capaz de hibridar con dicha secuencia de nucleótidos en condiciones de alta rigurosidad. El uso de dichas sondas en los productos desvelados en el presente documento en su preparación, y el uso de dichas sondas en los métodos descritos en el 35 presente documento son aspectos adicionales de la divulgación

En lo sucesivo, las referencias a "la secuencia de nucleótidos de SEQ ID NO X", también incluyen referencia a secuencias de nucleótidos capaces de hibridar en condiciones de alta rigurosidad con SEQ ID NO X a no ser que el contexto dicte otra cosa.

40 Los oligonucleótidos de este conjunto de sondas puede variar de tamaño dependiendo de qué secuencia de nucleótidos comprendan. En general, los oligonucleótidos pueden comprender hasta 100 nucleótidos, preferentemente hasta 80, 60, 50, 40, 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11 o hasta 10 nucleótidos. Los oligonucleótidos del conjunto de sondas pueden 45 comprender al menos 9, preferentemente al menos 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 60 o al menos 80 nucleótidos. Los oligonucleótidos del conjunto de sondas pueden comprender al menos 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 40, 60 o al menos 80 nucleótidos además del número de nucleótidos en cualquier secuencia de SEQ ID NO 1 a 52 que esté presente en el oligonucleótido.

50 Los nucleótidos de los oligonucleótidos del conjunto de sondas pueden ser de cualquier tipo de nucleótido siempre que la especificidad o eficacia de hibridación y, si es necesario, la eficacia de polimerización de ácido nucleico o eficacia de amplificación de ácido nucleico dependiente de cebador no se ve afectada perjudicialmente. Los oligonucleótidos pueden por lo tanto ser desoxirribonucleótidos, ribonucleótidos, modificaciones de los mismos (por 55 ejemplo PNA, morfolino-, LNA) y mezclas de los mismos. Se prefieren oligonucleótidos de ADN y oligonucleótidos de ADN modificados por LNA.

Los nucleótidos correspondientes a SEQ ID NO 1 a 52 pueden hallarse en cualquier parte de las sondas oligonucleotídicas siempre que los oligonucleótidos puedan hibridar con la secuencia diana complementaria a la 60 SEQ NO en consideración y, si se requiere, pueden efectuar una reacción de extensión ácido nucleico. En algunas realizaciones el nucleótido 3' de cualquier secuencia de SEQ ID NO 1 a 52 que esté presente en el oligonucleótido es el nucleótido 3' del oligonucleótido.

En otros aspectos los oligonucleótidos consisten esencialmente en una secuencia seleccionada de SEQ ID NO 1 a 52. Por lo tanto, los oligonucleótidos tendrán una secuencia de nucleótidos seleccionada de SEQ ID NO 1 a 52 y 1, 65

2, 3, 4 o 5 nucleótidos adicionales. En otros aspectos los oligonucleótidos consistirán en una secuencia seleccionada de SEQ ID NO 1 a 52.

A no ser que se indique de otro modo, o se dicte por el contexto específico, todas las secuencias de nucleótidos se enumeran en el presente documento de 5' a 3' en línea con la convención en este campo técnico.

Las condiciones de alta rigurosidad para hibridación se definen como SSC 2x/formamida al 50 % a 50 °C para condiciones de unión y SCC 2 x a 65 °C para condiciones de lavado (en las que SSC = NaCl 0,15 M, citrato sódico 0,015 M, pH 7,2).

Las secuencias de nucleótidos que pueden hibridar con una de SEQ ID NO 1 a 52 en condiciones de alta rigurosidad hibridarán con todos, o sustancialmente todos, los nucleótidos en las secuencias de SEQ ID NO 1 a 52, por ejemplo una serie de nucleótidos contiguos con varios nucleótidos que suman hasta al menos 50 % preferentemente al menos 55, 60, 65, 70, 75, 80, 85, 90 o 95 % del número total de nucleótidos en la secuencia de la SEQ ID NO que se considera.

Visto de forma alternativa, las secuencias de nucleótidos que pueden hibridar con las secuencias de nucleótidos de una de las SEQ ID NO 1 a 26 o 27 a 52 en condiciones de alta rigurosidad pueden ser las secuencias de nucleótidos que corresponden a la secuencia de nucleótidos de SEQ ID NO 27 a 52 o 1 a 26, respectivamente pero con hasta 40 % de las bases (adenina, timina/uracilo, guanina o citosina) en las secuencias de nucleótidos de SEQ ID NO 27 a 52 o 1 a 26, que se sustituyen con una base diferente. Preferentemente se sustituirán hasta 35, 30, 25, 20, 15, 10 o 5 % de las bases. Dicho de otro modo, las secuencias de nucleótidos que pueden hibridar con las secuencias de nucleótidos de una de las SEQ ID NO 1 a 26 o 27 a 52 en condiciones de alta rigurosidad pueden ser las secuencias de nucleótidos que corresponden a la secuencia de nucleótidos de SEQ ID NO 27 a 52 o 1 a 26, respectivamente pero con hasta 5, 4, 3 o 2 bases sustituidas o solamente una única sustitución de base. La base que se sustituye en la secuencia puede ser cualquier base convencional o no convencional, de origen natural o sintética.

Las secuencias de nucleótidos que pueden hibridar con las SEQ ID NO 1 a 26 o 27 a 52 en condiciones de alta rigurosidad serán preferentemente de 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27 o 28 nucleótidos de longitud, y consisten en una parte contigua de las secuencias de nucleótidos de SEQ ID 27 a 52 o 1 a 26, respectivamente, con las sustituciones anteriormente descritas.

Preferentemente la sustitución o las sustituciones de bases suceden en o cerca del extremo 5' de las secuencia de nucleótidos, por ejemplo en los últimos 15, 10 o 5 nucleótidos 5' en la secuencia. Dicho de otro modo, la sustitución o las sustituciones de bases preferentemente no suceden en o cerca del extremo 3' de la secuencia de nucleótidos, por ejemplo en los 2, 3, 4, 5, 10 o 15 nucleótidos 3'. En otras realizaciones el nucleótido 3' no tendrá una base sustituida.

En ciertos aspectos cualquiera de los oligonucleótidos del conjunto de sondas que comprende una secuencia de nucleótidos de una cualquiera de SEQ ID NO 1 a 26 puede tener un resto C inmediatamente 3' de dicha secuencia de nucleótidos o cualquiera de los oligonucleótidos que comprenden una secuencia de nucleótidos de una cualquiera de SEQ ID NO 27 a 52 puede tener un resto G inmediatamente 5' de dicha secuencia de nucleótidos.

Los oligonucleótidos del conjunto de sondas pueden marcarse con un resto para ayudar a la detección o manipulación. Se conocen en la técnica un gran número de restos adecuados y métodos de marcaje y se describen en la bibliografía. Muchos restos pueden realizar ambas funciones. Puede usarse cualquier molécula detectable o generadora de señal o molécula indicadora. Los marcadores convenientes incluyen marcadores colorimétricos, quimioluminiscentes, cromogénicos, radiactivos y fluorescentes, pero también pueden usarse métodos de marcaje enzimáticos (por ejemplo, colorimétricos, luminiscentes, cromogénicos) o basados en anticuerpos o sistemas generadores de señal. Por lo tanto el término "marcador" como se usa en el presente documento incluye no solamente restos pasivos o que proporcionan señal detectable directamente, sino también cualquier resto que genere una señal o tome parte en una reacción de generación de señal o que pueda detectarse indirectamente de alguna manera. Por ejemplo el resto puede ser biotina y la detección puede ser indirecta mediante estreptavidina que porta un resto colorimétrico, quimioluminiscente, cromogénico, radiactivo o fluorescente.

El marcador puede comprender, en algunas realizaciones, una pluralidad de restos que contribuye a la producción detectable general del marcador. Variando la identidad y/o las proporciones relativas de estos restos, puede construirse una amplia paleta de marcadores únicos. Por ejemplo, puede usarse una pluralidad de colorantes, por ejemplo luminiscentes (por ejemplo bioluminiscentes, quimioluminiscentes, fotoluminiscentes, radioluminiscentes, sonoluminiscentes, etc.) que combinan para proporcionar una identificación espectral electromagnética única tras la excitación. Variando las proporciones de los colorantes seleccionados puede conseguirse diferenciación adicional en la identificación espectral. También se prevén identificaciones basadas en la absorción de ciertas longitudes de onda de radiación electromagnética.

La fluoresceína u otros nucleótidos marcados con fluorescencia son particularmente adecuados para incorporación en los cebadores, y permiten la detección directamente por fluorescencia o indirectamente por interacciones de

anticuerpo. Estos están disponibles en el mercado. Los cebadores pueden marcarse por, por ejemplo, [³⁵S], [³H] o [³²P] como se describe en Syvänen, A.C. *et al.* Genomics 8, 684-692. Puede usarse cualquier resto de unión como un marcador, por ejemplo un fragmento de anticuerpo, marcador de His, biotina o estreptavidina. Estos pueden incorporarse en forma de nucleótidos marcados.

Algunos de o todos los oligonucleótidos del conjunto de sondas pueden proporcionarse inmovilizados en uno o más soportes sólidos para su uso en la invención. En otras realizaciones los oligonucleótidos del conjunto de sondas pueden inmovilizarse en uno o más soportes sólidos antes de su uso. Se unen copias individuales o preferentemente múltiples de las sondas oligonucleotídicas a dichos soportes sólidos, por ejemplo 10 o más, por ejemplo al menos 100 copias de cada sonda única están presentes.

Pueden asociarse una o más sondas oligonucleotídicas del conjunto de sondas, cada una de una cierta secuencia, con soportes sólidos separados que juntos forman un conjunto de sondas inmovilizado en múltiples soportes sólidos, por ejemplo una o más sondas oligonucleotídicas del conjunto de sondas pueden inmovilizarse en múltiples perlas, membranas, filtros, microplacas biológicas, etc. Los soportes sólidos de las diferentes partes del conjunto de sondas se asocian convenientemente de forma física aunque las señales asociadas con cada sonda (generadas como se describe en lo sucesivo en el presente documento) deben poder determinarse por separado.

Como alternativa, las sondas pueden inmovilizarse en partes discretas del mismo soporte sólido, por ejemplo cada sonda oligonucleotídica de una cierta secuencia, típicamente en múltiples copias, puede inmovilizarse en una parte definida y discreta o región de una única microplaca, placa, filtro o membrana, por ejemplo, para generar una matriz.

Puede usarse también una combinación de dichas técnicas, por ejemplo pueden usarse varios soportes sólidos que portan cada uno varias sondas de secuencia diferente inmovilizadas en los mismos.

La expresión "soporte sólido" significará cualquier material sólido capaz de unirse con oligonucleótidos, por ejemplo mediante interacción hidrófoba, iónica o covalente.

"Inmovilización" como se usa en el presente documento se refiere a asociación reversible o irreversible de las sondas con dicho soporte sólido. Si es reversible, las sondas permanecen asociadas con el soporte sólido durante un tiempo suficiente para llevar a cabo los métodos de la invención.

Se conocen en la técnica soportes de inmovilización adecuados con los que pueden unirse los oligonucleótidos e incluyen cualquiera de los soportes bien conocidos o matrices que se usan ampliamente en la actualidad o se han propuesto para inmovilización, separación, etc. de oligonucleótidos. Dichos materiales incluyen, pero sin limitación, cualquier polímero orgánico sintético tal como poliestireno, polivinilcloruro, polietileno o acetato de nitrocelulosa y celulosa; o agarosa, celulosa, alginato, teflón o látex; o unas superficies activadas con tosilo; o superficie de vidrio o nailon o cualquiera que porte un grupo adecuado para acoplamiento covalente de ácidos nucleicos. Estas pueden tomar la forma de partículas, láminas geles, filtros, membranas, fibras, capilares, microplacas o tiras de microtitulación, portaobjetos, tubos, placas o pocillos, etc. Se conocen en la técnica de modo similar métodos para inmovilizar o unir oligonucleótidos con los soportes sólidos. Se prefieren particularmente microplacas de ADN (microplacas, placas de vidrio) habituales ahora en procedimientos de biología molecular. En otras realizaciones pueden usarse tiras de membrana sobre las que pueden aplicarse puntualmente los oligonucleótidos y después reticularse por UV. Como alternativa, puede realizarse unión indirectamente mediante el uso de un resto de unión portado en las sondas oligonucleotídicas y/o soporte sólido. Por lo tanto, por ejemplo, puede usarse un par de compañeros de unión de afinidad, tales como avidina, estreptavidina o biotina, ADN o proteína de unión a ADN (por ejemplo la proteína represora de lac I o la secuencia operadora de lac con la que se une), anticuerpos (que pueden ser mono o policlonales), fragmentos de anticuerpos o los epítopos o haptenos de anticuerpos. En estos casos, se une un compañero de unión del par de unión con (o es inherentemente parte de) el soporte sólido y el otro compañero se une con (o es inherentemente parte de) las moléculas de ácido nucleico.

Como se usa en el presente documento un "par de unión de afinidad" se refiere a dos componentes que reconocen y se unen entre sí específicamente (es decir en preferencia con respecto a unión con otras moléculas).

Puede realizarse unión de grupos funcionales apropiados con el soporte sólido por métodos bien conocidos en la técnica, que incluyen por ejemplo unión mediante grupos hidroxilo, carboxilo, aldehído o amino que pueden proporcionarse tratando el soporte sólido para proporcionar recubrimientos de superficie adecuados. Pueden producirse soportes sólidos que presentan restos apropiados para la unión del compañero de unión mediante métodos rutinarios conocidos en la técnica.

Puede realizarse unión de grupos funcionales apropiados con las sondas oligonucleotídicas de la invención mediante ligamiento o introducirse durante síntesis o amplificación, por ejemplo usando cebadores que portan un resto apropiado, tal como biotina o una secuencia particular para captura.

Cada sonda oligonucleotídica de una cierta secuencia puede asociarse con un soporte sólido separado, por ejemplo una perla o una microesfera, que tiene un marcador particular de modo que se forme una población, o pluralidad de

poblaciones, de partículas que tienen el mismo marcador y la misma sonda inmovilizada en las mismas. La detección de un acontecimiento de hibridación que sucede en una partícula con un marcador particular proporcionará información sobre la secuencia de la sonda implicada en ese acontecimiento.

5 Las partículas pueden marcarse de cualquier manera conveniente, por ejemplo usando uno o más de los marcadores descritos anteriormente. En una realización el marcador de partícula no estará o comprenderá un oligonucleótido, o un ácido nucleico, o un oligonucleótido marcado o ácido nucleico marcado. Convenientemente el soporte sólido en partículas de estas realizaciones se marcará con un colorante, por ejemplo un colorante
10 luminiscente (por ejemplo bioluminiscente, quimioluminiscente, fotoluminiscente, radioluminiscente, sonoluminiscente, etc.), o una pluralidad de marcadores (o proporciones de los mismos) que se combinan para proporcionar una identificación espectral electromagnética única tras la excitación. También se prevén identificaciones basadas en la absorción de ciertas longitudes de onda de radiación electromagnética.

15 Convenientemente el colorante será fluorescente, por ejemplo comprenderá fluoróforos rojos o infrarrojos, por ejemplo ficoeritrina.

El marcador puede inmovilizarse sobre y/o en la partícula, por ejemplo mediante unión covalente directa con el sustrato de la partícula o puede unirse con otra molécula que a su vez se inmoviliza sobre y/o en la partícula. El
20 marcador también puede incorporarse en y/o sobre la partícula por medios no covalentes, por ejemplo mediante inmovilización, absorción o adsorción de las moléculas que componen el marcador en o sobre un sustrato de la partícula, o mediante inmovilización en vacío o vacíos dentro del sustrato y/o su superficie.

Como alternativa, la partícula comprende nanopartículas sobre las que y/o en las que el marcador se ha inmovilizado o incorporado.

25 El marcador puede aplicarse a la partícula después de producirse, o el marcador puede incorporarse o inmovilizarse en y/o sobre la partícula durante su producción, por ejemplo durante la reticulación de un sustrato polimérico.

Preferentemente el marcador de la sonda o las sondas será distinguible del marcador de la partícula o partículas. En
30 realizaciones preferidas el marcador de las partículas será detectable al mismo tiempo que el marcador de la sonda o las sondas. Preferentemente las partículas marcadas también serán magnéticas, por ejemplo, paramagnéticas o superparamagnéticas.

Se fabrican soportes sólidos en partículas adecuados por Luminex Corp. Véase por ejemplo los documentos
35 WO01/13120, WO01/13119, WO97/14028 y WO99/19515.

Se proporcionan partículas adicionales que pueden usarse en el trabajo de la invención en los documentos
40 US4267234, US4267235, US4552812, US4677138, US5194300, US4774189, US5073498, US4717655, US5723218, US5326692, US5716855, US5573909 y US5786219. Otros soportes sólidos adecuados se fabrican por Illumina, Inc. Véase por ejemplo los documentos WO00/39587, WO 01/18524, WO01/59432 y WO02/00336.

Preferentemente el soporte es magnético (preferentemente paramagnético o superparamagnético) por ejemplo
partículas magnéticas, por ejemplo perlas magnéticas.

45 Como alternativa, no se usa ninguno de los oligonucleótidos en forma inmovilizada, y en general esto se prefiere.

Preferentemente el conjunto de sondas oligonucleotídicas definido en el presente documento comprende:

(a) un oligonucleótido que comprende una secuencia de nucleótidos seleccionada de SEQ ID NO 1 o una secuencia
50 capaz de hibridar con SEQ ID NO 27 en condiciones de alta rigurosidad;
(b) un oligonucleótido que comprende una secuencia de nucleótidos seleccionada de SEQ ID NO 2 o una secuencia capaz de hibridar con SEQ ID NO 28 en condiciones de alta rigurosidad;
(c) un oligonucleótido que comprende una secuencia de nucleótidos seleccionada de SEQ ID NO 3 o una secuencia
55 capaz de hibridar con SEQ ID NO 29 en condiciones de alta rigurosidad;
(d) un oligonucleótido que comprende una secuencia de nucleótidos seleccionada de SEQ ID NO 4 o una secuencia capaz de hibridar con SEQ ID NO 30 en condiciones de alta rigurosidad; y
(e) uno o más oligonucleótidos que comprenden una secuencia de nucleótidos seleccionada de SEQ ID NO 5 a 8 o
una secuencia capaz de hibridar con SEQ ID NO 31 a 34 en condiciones de alta rigurosidad; y opcionalmente
(f) uno o más oligonucleótidos que comprenden una secuencia de nucleótidos seleccionada de SEQ ID NO 9 a 18,
60 una secuencia capaz de hibridar con SEQ ID NO 35 a 44 en condiciones de alta rigurosidad, una secuencia de nucleótidos seleccionada de SEQ ID NO 19 a 26 o una secuencia capaz de hibridar con SEQ ID NO 45 a 52 en condiciones de alta rigurosidad. Preferentemente, en esta realización no se inmoviliza ninguna de las sondas oligonucleotídicas en un soporte sólido.

65 En otros aspectos preferidos el conjunto de sondas oligonucleotídicas definido en el presente documento comprende:

(a) un oligonucleótido que comprende una secuencia de nucleótidos seleccionada de SEQ ID NO 27 o una secuencia capaz de hibridar con SEQ ID NO 1 en condiciones de alta rigurosidad;
 (b) un oligonucleótido que comprende una secuencia de nucleótidos seleccionada de SEQ ID NO 28 o una secuencia capaz de hibridar con SEQ ID NO 2 en condiciones de alta rigurosidad;
 5 (c) un oligonucleótido que comprende una secuencia de nucleótidos seleccionada de SEQ ID NO 29 o una secuencia capaz de hibridar con SEQ ID NO 3 en condiciones de alta rigurosidad;
 (d) un oligonucleótido que comprende una secuencia de nucleótidos seleccionada de SEQ ID NO 30 o una secuencia capaz de hibridar con SEQ ID NO 4 en condiciones de alta rigurosidad, y
 10 (e) uno o más oligonucleótidos que comprenden una secuencia de nucleótidos seleccionada de SEQ ID NO 31 a 34 o una secuencia capaz de hibridar con SEQ ID NO 5 a 8 en condiciones de alta rigurosidad; y opcionalmente
 (f) uno o más oligonucleótidos que comprenden una secuencia de nucleótidos seleccionada de SEQ ID NO 35 a 44, una secuencia capaz de hibridar con SEQ ID NO 9 a 18 en condiciones de alta rigurosidad, SEQ ID NO 45 a 52 o una secuencia capaz de hibridar con SEQ ID NO 19 a 26 en condiciones de alta rigurosidad. Preferentemente, en esta realización, todas las sondas oligonucleotídicas se inmovilizan en un soporte sólido, o una pluralidad de
 15 soportes sólidos.

En un aspecto preferido adicional más, el conjunto de sondas oligonucleotídicas es una combinación de los dos aspectos preferidos precedentes, en particular en la que el conjunto de sondas preferido anterior no se inmoviliza en un soporte sólido. En otras palabras, se proporciona un kit que comprende, en dos partes discretas, los dos
 20 conjuntos de sondas preferidos definidos anteriormente.

El conjunto de sondas oligonucleotídicas de la invención se define por las reivindicaciones 1-7. Las sondas oligonucleotídicas del conjunto de sondas son capaces de hibridar con secuencias en ARNr 16S bacteriano y ADNr que son específicos para bacterias particulares o grupos de bacterias. SEQ ID NO 1 y 27 se diseñan para dirigirse a secuencias en el ARNr 16S y ADNr de proteobacterias. SEQ ID NO 2 y 28 se diseñan para dirigirse a secuencias en
 25 ARNr 16S y ADNr de *Firmicutes* (*Lactobacillales*, *Clostridium perfringens*, *Staphylococcus*). SEQ ID NO 3 y 29 se diseñan para dirigirse a secuencias en el ARNr 16S y ADNr de *Firmicutes* (*Clostridia*, *Bacillales*, *Enterococcus*, *Lactobacillus*). SEQ ID NO 4 y 30 se diseñan para dirigirse a secuencias en el ARNr 16S y ADNr de actinobacterias. Los organismos que presentan las secuencias diana de SEQ ID NO 5 a 26 y 31 a 52 se enumeran en la Tablas 1, 2 y 3. Estas bacterias son bacterias halladas típicamente en el tracto GI.

El conjunto de sondas oligonucleotídicas puede usarse por lo tanto para analizar muestras del tracto GI y proporcionar un perfil de la microbiota del tracto GI.

Por lo tanto en otro aspecto la invención proporciona un método para realizar el perfil de la microbiota del tracto GI de un sujeto, comprendiendo dicho método:

(i) poner en contacto la muestra del tracto GI de dicho sujeto con un conjunto de sondas oligonucleotídicas como se define en las reivindicaciones 1-7;
 40 (ii) someter la muestra y el conjunto de sondas a condiciones que permiten la hibridación de las sondas con sus secuencias de nucleótidos diana dentro de moléculas de ácido nucleico en dicha muestra; y
 (iii) para cada oligonucleótido en dicho conjunto de sondas, determinar la cantidad de su secuencia de nucleótidos diana que está presente en dicha muestra.

El perfil de la microbiota del tracto GI del sujeto puede prepararse después a partir de las cantidades relativas de dicha secuencia diana.

En este aspecto de la invención las secuencias de nucleótidos diana para SEQ ID NO 1 a 26 son las secuencias completamente complementarias de las mismas, es decir SEQ ID NO 27 a 52, respectivamente. De forma similar, las secuencias de nucleótidos diana para SEQ ID NO 27 a 52 son las secuencias completamente complementarias de las mismas, es decir SEQ ID NO 1 a 26, respectivamente. En ciertas realizaciones las secuencias de nucleótidos diana para SEQ ID NO 1 a 26 son las secuencias de SEQ ID NO 27 a 52, respectivamente, con un resto G inmediatamente 5' de dicha secuencia de nucleótidos. De forma similar, las secuencias de nucleótidos diana para SEQ ID NO 27 a 52 son las secuencias de SEQ ID NO 1 a 26, respectivamente con un resto C inmediatamente 3' de dicha secuencia de nucleótidos.
 55

La cantidad de secuencia diana puede determinarse por cualquier medio conveniente y el experto en la materia estará familiarizado con muchos de dichos medios. Esta puede ser una medición parcialmente, semi o completamente cuantitativa, pero también puede ser una medida cualitativa (o relativa) en la que los resultados para cada diana se comparan simplemente entre sí sin fijarse valores numéricos. Como se analiza posteriormente, en algunas realizaciones se realiza medición cuantitativa y los datos obtenidos se analizan con técnicas estadísticas para determinar las características estadísticamente significativas del perfil de microbiota.

En una realización la cantidad de cada secuencia diana se determina usando los oligonucleótidos del conjunto de sondas de la invención con marcadores unidos a los mismos que permitirán la detección por medios directos o
 65 medios indirectos. En otras palabras las sondas oligonucleotídicas de la invención se usan sencillamente como

sondas oligonucleotídicas convencionales. Se han descrito anteriormente marcadores adecuados. Después del contacto de dichas sondas con la muestra, en condiciones que permiten la hibridación de las sondas con sus secuencias diana, y típicamente después de una etapa (o etapas) para retirar oligonucleótidos marcados no unidos y/u oligonucleótidos no específicamente, la fuerza de la señal del marcador de cada sonda que surge de la muestra

- 5 que se investiga (es decir la cantidad de marcador unido a la muestra) será proporcional a la cantidad de oligonucleótido hibridado, y por tanto su secuencia diana. En realizaciones preferidas el marcador se selecciona de modo que sea detectable solamente cuando la sonda esté hibridada con su diana. En dichas realizaciones, se reduce la necesidad de retirar la sonda no unida.
- 10 Puede usarse cualquier medio conveniente para retirar cualquier sonda no unida o unida no específicamente, por ejemplo con una o más etapas de lavado (por ejemplo con agua o una solución tamponada que puede contener formamida y/o un detergente), electroforesis, centrifugación, captura en soportes sólidos, cromatografía o cualquier combinación de los mismos. Se han descrito anteriormente soportes sólidos adecuados. En otra realización las sondas pueden portar un resto de unión, o el marcador puede ser un resto de unión, que permitirá la manipulación
- 15 de las sondas y cualquier parte de la muestra hibridada con las mismas. Se han analizado anteriormente restos de unión adecuados.

Por lo tanto, la invención proporciona un método para realizar el perfil de la microbiota del tracto GI de un sujeto, comprendiendo dicho método:

- 20 (i) poner en contacto la muestra del tracto GI de dicho sujeto con un conjunto de sondas oligonucleotídicas como se define en las reivindicaciones 1-5, en el que cada oligonucleótido tiene un marcador unido al mismo;
- (ii) someter la muestra y el conjunto de sondas a condiciones que permiten la hibridación de las sondas con sus secuencias de nucleótidos diana dentro de moléculas de ácido nucleico en dicha muestra; y
- 25 (iii) para cada oligonucleótido marcado en dicho conjunto de sondas, determinar la cantidad de dicho marcador unido a dicha muestra determinando la fuerza de la señal del marcador que surge de la muestra. La cantidad de marcador unido a la muestra es indicativa de la cantidad de la secuencia diana para ese oligonucleótido marcado en dicha muestra.
- 30 En una realización preferida el método comprenderá una etapa entre las etapas (i) y (ii) en la que se retira oligonucleótido no unido y/u oligonucleótido unido de forma no específica.

En otra realización la cantidad de cada secuencia de nucleótidos diana presente en la muestra se determina usando un conjunto de sondas oligonucleotídicas de la invención, en particular los conjuntos que comprenden sondas oligonucleotídicas que consisten en secuencias de nucleótidos seleccionadas de SEQ ID NO 1 a 26, como un

35 conjunto de sondas que se marcan solamente cuando se hibridan con sus secuencias diana. En algunas realizaciones los oligonucleótidos del conjunto de sondas pueden portar ya un marcador que es diferente del marcador usado para marcar selectivamente las sondas. La fuerza de la señal de las sondas marcadas selectivamente que surgen de la muestra que se investiga (es decir la cantidad de marcadores unidos a la muestra) será proporcional a la cantidad de oligonucleótido hibridado y a su vez la cantidad de secuencia diana.

40

Como se ha mencionado previamente, dependiendo de las condiciones empleadas, esta puede ser una medición parcialmente, semi o completamente cuantitativa, pero también puede ser una medida cualitativa (o relativa) en la que los resultados para cada secuencia diana se comparan sencillamente entre sí sin fijarse valores numéricos.

- 45 Convenientemente, puede conseguirse marcaje selectivo usando nucleótidos marcados, es decir mediante incorporación en la sonda oligonucleotídica de un nucleótido que porta un marcador. En otras palabras, puede producirse marcaje selectivo mediante extensión de cadena de la sonda oligonucleotídica usando una enzima polimerasa que incorpora un nucleótido marcado, preferentemente un didesoxinucleótido marcado (por ejemplo ddATP, ddCTP, ddGTP, ddTTP, ddUTP) más preferentemente ddCTP marcado, más preferentemente un ddCTP marcado con fluorescencia, por ejemplo marcado con TAMRA, ddCTP o un ddCTP marcado con biotina. Este enfoque para la detección de secuencias de nucleótidos específicas se denomina en ocasiones análisis de extensión de cebadores. Los expertos en la materia conocen bien técnicas de análisis de extensión de cebadores adecuadas, por ejemplo las técnicas desveladas en el documento WO99/50448. Se han descrito anteriormente marcadores
- 50 adecuados. Se mencionan en particular marcadores fluorescentes y biotina.
- 55

En el caso de sondas oligonucleotídicas que terminan con SEQ ID NO 1 a 26 en su 3', el marcador será preferentemente un ddCTP marcado, por ejemplo un ddCTP marcado con TAMRA o biotina. Más preferentemente en esta realización el conjunto de sondas de la invención comprenderá oligonucleótidos que consisten en SEQ ID

60 NO 1 a 26 y el marcador será un ddCTP marcado, por ejemplo un ddCTP marcado con TAMRA o biotina.

La detección de las sondas marcadas puede ser por cualquier medio conveniente para el marcador que se usa. El experto en la materia sería capaz de idear métodos adecuados basados en su selección de marcadores. En realizaciones preferidas, los marcadores son marcadores fluorescentes (por ejemplo TAMRA) y en dichas realizaciones las sondas marcadas con fluorescencia pueden detectarse y, si se requiere, cuantificarse usando un dispositivo que puede medir la intensidad (o fuerza) de señales fluorescentes. Un marcador de biotina puede

65

detectarse indirectamente exponiendo el marcador a estreptavidina, u otra molécula de unión a biotina, que porta un resto detectable, por ejemplo un marcador colorimétrico, quimioluminiscente, cromogénico, radiactivo o fluorescente. En algunas realizaciones, se producirá detección después de que las sondas marcadas hayan experimentado manipulación para retirar, al menos parcialmente, contaminantes (por ejemplo sondas no marcadas, marcador en exceso y otros reactivos usados en la reacción de marcaje). De nuevo, el experto en la materia estará muy familiarizado con técnicas que pueden conseguir esto, como ejemplo se menciona electroforesis (por ejemplo electroforesis en gel, por ejemplo gel capilar), centrifugación, cromatografía y técnicas basadas en filtración, captura en soportes sólidos, o cualquier combinación de las mismas.

En otras realizaciones preferidas las sondas oligonucleotídicas marcadas selectivamente se detectan después de una etapa en la que las sondas oligonucleotídicas de la etapa de marcaje selectivo (es decir marcadas y no marcadas) o las sondas oligonucleotídicas marcadas selectivamente solamente, se hibridan con secuencias de nucleótidos que son parcialmente, o preferentemente completamente, complementarias de las sondas oligonucleotídicas.

Convenientemente, las secuencias de nucleótidos complementarias pueden proporcionarse en uno o más soportes sólidos, por ejemplo los descritos previamente.

En particular, el conjunto de sondas oligonucleotídicas que experimenta marcaje selectivo comprende:

- (a) un oligonucleótido que comprende una secuencia de nucleótidos seleccionada de SEQ ID NO 1 o una secuencia capaz de hibridar con SEQ ID NO 27 en condiciones de alta rigurosidad;
- (b) un oligonucleótido que comprende una secuencia de nucleótidos seleccionada de SEQ ID NO 2 o una secuencia capaz de hibridar con SEQ ID NO 28 en condiciones de alta rigurosidad;
- (c) un oligonucleótido que comprende una secuencia de nucleótidos seleccionada de SEQ ID NO 3 o una secuencia capaz de hibridar con SEQ ID NO 29 en condiciones de alta rigurosidad;
- (d) un oligonucleótido que comprende una secuencia de nucleótidos seleccionada de SEQ ID NO 4 o una secuencia capaz de hibridar con SEQ ID NO 30 en condiciones de alta rigurosidad; y
- (e) uno o más oligonucleótidos que comprenden una secuencia de nucleótidos seleccionada de SEQ ID NO 5 a 8 o una secuencia capaz de hibridar con SEQ ID NO 31 a 34 en condiciones de alta rigurosidad; y opcionalmente
- (f) uno o más oligonucleótidos que comprenden una secuencia de nucleótidos seleccionada de SEQ ID NO 9 a 18, una secuencia capaz de hibridar con SEQ ID NO 35 a 44 en condiciones de alta rigurosidad, SEQ ID NO 19 a 26 o una secuencia capaz de hibridar con SEQ ID NO 45 a 52 en condiciones de alta rigurosidad.

Después del marcaje selectivo del conjunto de sondas anterior, las sondas marcadas selectivamente se aplican a un conjunto de sondas unido a soporte sólido que comprende:

- (a) un oligonucleótido que comprende una secuencia de nucleótidos seleccionada de SEQ ID NO 27 o una secuencia capaz de hibridar con SEQ ID NO 1 en condiciones de alta rigurosidad;
- (b) un oligonucleótido que comprende una secuencia de nucleótidos seleccionada de SEQ ID NO 28 o una secuencia capaz de hibridar con SEQ ID NO 2 en condiciones de alta rigurosidad;
- (c) un oligonucleótido que comprende una secuencia de nucleótidos seleccionada de SEQ ID NO 29 o una secuencia capaz de hibridar con SEQ ID NO 3 en condiciones de alta rigurosidad;
- (d) un oligonucleótido que comprende una secuencia de nucleótidos seleccionada de SEQ ID NO 30 o una secuencia capaz de hibridar con SEQ ID NO 4 en condiciones de alta rigurosidad; y
- (e) uno o más oligonucleótidos que comprenden una secuencia de nucleótidos seleccionada de SEQ ID NO 31 a 34 o una secuencia capaz de hibridar con SEQ ID NO 5 a 8 en condiciones de alta rigurosidad; y opcionalmente
- (f) uno o más oligonucleótidos que comprenden una secuencia de nucleótidos seleccionada de SEQ ID NO 35 a 44, una secuencia capaz de hibridar con SEQ ID NO 9 a 18 en condiciones de alta rigurosidad, SEQ ID NO 45 a 52 o una secuencia capaz de hibridar con SEQ ID NO 19 a 26 en condiciones de alta rigurosidad. Preferentemente, en esta realización, todas las sondas oligonucleotídicas se inmovilizan en uno o más soportes sólidos.

Preferentemente, se seleccionan componentes (e) y (f) de cada conjunto de sondas para corresponder entre sí (es decir cada sonda en los componentes (e) y (f) de cada conjunto de sondas tiene una secuencia complementaria en componentes (e) y (f) del otro conjunto de sondas).

La invención proporciona por lo tanto un método para realizar el perfil de la microbiota del tracto GI de un sujeto, comprendiendo dicho método:

- (i) poner en contacto la muestra del tracto GI de dicho sujeto con un conjunto de sondas oligonucleotídicas como se define en las reivindicaciones 1-5;
- (ii) someter la muestra y el conjunto de sondas a condiciones que permiten la hibridación de las sondas con sus secuencias de nucleótidos diana dentro de moléculas de ácido nucleico en dicha muestra;
- (iii) marcar selectivamente las sondas oligonucleotídicas del conjunto de sondas cuando se hibriden con su secuencia de nucleótidos diana; y
- (iv) determinar la cantidad de cada sonda oligonucleotídica marcada producida en la etapa (iii).

La cantidad de cada sonda oligonucleotídica marcada es indicativa de la cantidad de la secuencia diana para ese oligonucleótido marcado en dicha muestra.

5 En algunas realizaciones, la etapa (iv) comprende hibridación de las sondas oligonucleotídicas de la etapa de marcaje con secuencias de nucleótidos complementarias de esos oligonucleótidos.

10 En una realización adicional la cantidad de cada secuencia de nucleótidos diana presente en la muestra se determina por marcaje de los ácidos nucleicos en la muestra antes de la etapa de poner en contacto la muestra con el conjunto de sondas oligonucleotídicas de la invención. Simplemente evaluando la cantidad de hibridación de ácido nucleico marcado con las sondas del conjunto de sondas puede determinarse la cantidad de la secuencia de nucleótidos diana para cada sonda oligonucleotídica que está presente en dicha muestra. En estas realizaciones se marca ARNr 16S o ADNr 16S bacteriano, particularmente las regiones que contienen las secuencias diana para los oligonucleótidos del conjunto de sondas, o ácidos nucleicos que comprenden dichas secuencias antes del contacto con el conjunto de sondas. Se han analizado anteriormente marcadores adecuados. Convenientemente el marcaje se produce cuando los ácidos nucleicos en la muestra se amplifican y/o transcriben de forma inversa antes del contacto con el conjunto de sondas como se analiza en más detalle posteriormente. Convenientemente, los ácidos nucleicos se marcan por incorporación de nucleótidos marcados durante una reacción de amplificación de ácido nucleico y/o una reacción de transcripción inversa.

20 En realizaciones adicionales tanto los oligonucleótidos del conjunto de sondas como los ácidos nucleicos en la muestra como se ha descrito anteriormente se marcan con restos que proporcionan una señal solamente cuando están en proximidad estrecha, por ejemplo cuando las sondas se hibridan con sus secuencias diana en el ácido nucleico.

25 En otra realización la cantidad de cada secuencia de nucleótidos diana presente en la muestra se determina usando los oligonucleótidos del conjunto de sondas de la invención como cebadores en una o más reacciones de amplificación de ácido nucleico, por ejemplo una reacción de amplificación múltiple. Si se seleccionan las condiciones apropiadas, dicha reacción puede realizarse de modo que la cantidad de producto de amplificación obtenido para cada oligonucleótido del conjunto de sondas sea proporcional a la cantidad de cada secuencia de nucleótidos diana presente en la muestra. Por lo tanto, la cantidad de producto que la reacción de amplificación proporciona para cada oligonucleótido del conjunto de sondas es una medida de la cantidad de ese oligonucleótido que hibrida con la muestra del sujeto que se investiga y es a su vez proporcional a la cantidad de secuencia diana para ese oligonucleótido en la muestra y por tanto es proporcional a la cantidad de bacterias a las que se ha diseñado que se dirija el oligonucleótido en la muestra. En consecuencia, la cantidad de producto de amplificación puede usarse para determinar los niveles de estas bacterias en el tracto GI de un sujeto.

40 Como se ha mencionado anteriormente, dependiendo de las condiciones empleadas, esta puede ser una medición parcialmente, semi o completamente cuantitativa, pero también puede ser una medida cualitativa (o relativa) en la que los resultados para cada secuencia diana se comparan sencillamente entre sí sin fijarse valores numéricos.

45 Puede conseguirse amplificación por cualquier reacción de amplificación de ácido nucleico dependiente de cebador conveniente. Más convenientemente se usará la reacción en cadena de la polimerasa (PCR), aunque el experto en la materia conocería otras técnicas. Por ejemplo, puede usarse LAR/LCR, SDA, amplificación isotérmica mediada por bucle y amplificación basada en secuencias de ácido nucleico (NASBA)/3SR (replicación de secuencia autosostenible).

50 Se han desarrollado muchas variaciones de PCR, por ejemplo PCR en tiempo real (también conocida como PCR cuantitativa, qPCR), PCR de arranque en caliente, PCR competitiva, y así sucesivamente, y todas estas pueden emplearse según sea apropiado para las necesidades del experto.

55 En una realización básica de la invención usando una amplificación basada en PCR los oligonucleótidos del conjunto de sondas de la invención se ponen en contacto con una mezcla de reacción que contiene la muestra, un conjunto adecuado de segundos cebadores para formar un conjunto de pares de cebadores de trabajo y nucleótidos libres en un tampón adecuado en condiciones que permiten la hibridación. La termociclación de la mezcla resultante en presencia de una ADN polimerasa da como resultado amplificación de las secuencias diana para cada oligonucleótido, es decir secuencias características de las bacterias a las que se ha diseñado que se dirijan los oligonucleótidos del conjunto de sondas de la invención.

60 El rendimiento óptimo del proceso de PCR está influido por la elección de temperatura, el tiempo a esa temperatura, y la longitud del tiempo entre temperaturas para cada etapa en el ciclo. Un perfil de ciclación típico para amplificación por PCR es (a) 15 minutos de fusión de ADN a 95 °C; (b) 30 segundos de hibridación de cebadores a 50-65 °C; (c) 90 segundos de extensión de cebadores a 68-72 °C; (d) 30 segundos fusión de ADN a 95 °C; y las etapas (b)-(d) se repiten tantas veces como sea necesario para obtener el nivel deseado de amplificación.

65 Se han desarrollado modificaciones del método de PCR básico tal como qPCR (PCR en tiempo real) que pueden proporcionar información cuantitativa sobre el molde que se amplifica. Se han elaborado varios enfoques aunque las

dos técnicas más comunes usan colorantes fluorescentes de unión a ADN bicatenario o sondas indicadoras fluorescentes selectivas.

Los colorantes fluorescentes de unión a ADN bicatenario, por ejemplo SYBR Green, se asocian con el producto de amplificación como se produce y cuando se asocia en el colorante fluoresce. En consecuencia, midiendo la fluorescencia después de cada ciclo de PCR, puede supervisarse en tiempo real la cantidad relativa de producto de amplificación. Mediante el uso de patrones y controles internos, esta información puede traducirse en datos cuantitativos sobre la cantidad de molde al comienzo de la reacción.

Las sondas indicadoras fluorescentes usadas en qPCR son oligonucleótidos específicos de secuencia, típicamente ARN o ADN, que tienen una molécula indicadora fluorescente en un extremo y una molécula interruptora en el otro (por ejemplo la molécula indicadora está en el extremo 5' y una molécula interruptora en el extremo 3' o viceversa). La sonda se diseña de modo que el indicador se inactiva por el interruptor. La sonda también se diseña para hibridar selectivamente con regiones particulares de secuencia complementaria que podrían estar en el molde. Si estas regiones están entre los cebadores de PCR hibridados la polimerasa, si tiene actividad exonucleasa, degradará (despolimerizará) la sonda unida a medida que extiende la cadena ácido nucleico naciente que está polimerizando. Esto aliviará la inactivación y aumentará la fluorescencia. En consecuencia, midiendo la fluorescencia después de cada ciclo de PCR, la cantidad relativa de producto de amplificación puede supervisarse en tiempo real. Mediante el uso de patrones y controles internos, esta información puede traducirse a datos cuantitativos.

Por lo tanto, en otro aspecto la invención proporciona un método para realizar el perfil de la microbiota del tracto GI de un sujeto, comprendiendo dicho método:

- (i) poner en contacto la muestra del tracto GI de dicho sujeto con un conjunto de sondas oligonucleotídicas como se define en las reivindicaciones 1-5;
- (ii) someter la muestra y el conjunto de sondas a condiciones que permiten la hibridación de las sondas con sus secuencias de nucleótidos diana dentro de moléculas de ácido nucleico en dicha muestra;
- (iii) realizar una reacción de amplificación de ácidos nucleicos dependiente de cebador; y
- (iv) para cada oligonucleótido en el conjunto de sondas determinar la cantidad de producto de amplificación producido a partir del mismo en dicha reacción de amplificación de ácidos nucleicos dependiente de cebador.

La cantidad de dicho producto para cada oligonucleótido es indicativa de la cantidad de la secuencia diana para cada oligonucleótido en la muestra.

En una realización preferida la etapa (i) también comprenderá poner en contacto la muestra con un conjunto de oligonucleótidos que pueden actuar con el conjunto de oligonucleótidos de la invención en una reacción de amplificación de ácido nucleico, por ejemplo PCR, para producir un producto de amplificación para cada oligonucleótido del conjunto de sondas, suponiendo que está presente un molde adecuado en la muestra. En esta realización, cuando se empareja con un segundo conjunto de cebadores de amplificación adecuados, los oligonucleótidos que comprenden SEQ ID NO 1 a 26 actuarán como cebadores directos y los oligonucleótidos que comprenden SEQ ID NO 27 a 52 actuarán como cebadores inversos.

En esta realización el método puede implicar una pluralidad de amplificaciones de ácidos nucleicos dependientes de cebadores que se procesan en paralelo, implicando cada reacción una única sonda, o procesándose una o más amplificaciones de ácido nucleico dependientes de cebadores múltiples usándose dos o más sondas en la misma reacción.

Puede detectarse el producto de amplificación de cada oligonucleótido, y pueden determinarse las cantidades de producto de amplificación, por cualquier medio conveniente. En algún grado las técnicas factibles se dictarán por el número de oligonucleótidos del conjunto de sondas que se usan en cada reacción de amplificación (por ejemplo si la reacción es una reacción múltiple o no o el alcance de la multiplexación). El experto en la materia podría seleccionar técnicas apropiadas.

Se emplean rutinariamente un amplio número de técnicas como técnicas de laboratorio convencionales y en la bibliografía hay descripciones de enfoques más especializados. En su caso más simple la cantidad de producto de amplificación puede detectarse o determinarse por inspección visual de la mezcla de reacción al final de la reacción o en un punto temporal deseado. Típicamente el producto de amplificación se resolverá con ayuda de un marcador que puede unirse preferentemente con el producto de amplificación. Típicamente se usa una sustancia colorante, por ejemplo un colorante colorimétrico, fluorescente cromomérico o luminiscente (por ejemplo bromuro de etidio o SYBR green). En otras realizaciones se usa una sonda oligonucleotídica marcada que se une preferentemente con el producto de amplificación, en particular una sonda que se une preferentemente con sustancialmente todos los ácidos nucleicos amplificados individuales en el producto de amplificación. Una sonda adecuada podría basarse en la secuencia de nucleótidos de una o más de SEQ ID NO 1 a 52. Se han analizado anteriormente marcadores adecuados para la sonda. En algunas realizaciones la sonda puede proporcionarse en una forma no marcada sucediendo marcaje después de la unión preferente con el producto de amplificación, o unión preferente con sustancialmente todos los ácidos nucleicos amplificados individuales en el producto de amplificación.

Sin embargo, en algunos casos puede usarse simplemente un precipitante de ácido nucleico (por ejemplo sal y/o alcohol) para provocar que el producto de amplificación salga de la solución y sea visible sin marcaje.

Para ayudar a la visualización los componentes de producto de amplificación pueden dispersarse en o sobre un soporte sólido, por ejemplo por electroforesis (por ejemplo usando geles de agarosa o poliacrilamida), cromatografía (por ejemplo HPLC, TLC, afinidad, filtración en gel) o filtración, o una combinación de los mismos, antes o después del contacto con el marcador.

Dependiendo del marcador usado la detección puede hacerse más precisa usando tecnologías de detección ampliamente disponibles, por ejemplo películas sensibles a radiación y tecnologías de captura de imágenes digitales en combinación con análisis de imágenes asistido por ordenador, fotómetros, fluorímetros, colorímetros, contadores de centelleo y similares.

Preferentemente el producto de amplificación se separa del resto de la reacción de amplificación antes de ponerse en contacto con el marcador, por ejemplo en forma de una sonda oligonucleotídica marcada. Esto puede ser por cualquier medio conveniente, por ejemplo con una o más etapas de lavado (por ejemplo con agua o una solución tamponada que puede contener formamida y/o un detergente), electroforesis, centrifugación, captura en soportes sólidos de unión a ácido nucleico, cromatografía o cualquier combinación de los mismos. Convenientemente, la sonda puede proporcionarse en un soporte sólido efectuando de este modo separación del producto de amplificación del resto de la reacción de amplificación en una única etapa. En otra realización la sonda puede portar un resto de unión, o el marcador puede ser un resto de unión, que permita la manipulación de la sonda y cualquier producto de amplificación hibridado con la misma. Se han analizado anteriormente restos de unión adecuados.

Preferentemente se separará cualquier marcador no unido, por ejemplo en forma de una sonda oligonucleotídica marcada, del producto de amplificación antes de la etapa de detección. Esto puede ser por cualquier medio conveniente, por ejemplo con una o más etapas de lavado (por ejemplo con agua o una solución tamponada que puede contener formamida y/o un detergente), electroforesis, centrifugación, captura en soportes sólidos, cromatografía o cualquier combinación de los mismos. Se han descrito anteriormente soportes sólidos adecuados.

Si el método de amplificación usado es en sí mismo cuantitativo, por ejemplo métodos de amplificación en los que se incorporan patrones y controles internos (por ejemplo qPCR), el método de este aspecto de la invención también puede proporcionar datos cuantitativos. En estas realizaciones el método puede incluso fijar un valor numérico a la cantidad de secuencia diana presente en la muestra y por lo tanto la cantidad de las bacterias que contienen la secuencia diana en la muestra. Uno de dichos patrones internos sería amplificar una o más (por ejemplo al menos 2, 3, 5 o 10) muestras que tienen cantidades conocidas de la bacteria a la que se dirigen los oligonucleótidos del conjunto de sondas o de cantidades conocidas de secuencia diana en las mismas condiciones que la muestra de ensayo para proporcionar una curva patrón que representa la cantidad de producto de amplificación frente al número de organismos o cantidad de secuencia diana. La cantidad de producto de amplificación obtenida en la muestra de ensayo puede después traducirse a un valor numérico para la cantidad de secuencia diana y/o bacteria que contiene la secuencia diana de los oligonucleótidos del conjunto de sondas en la muestra.

En otras realizaciones, el progreso de la reacción de amplificación puede seguirse en tiempo real y el perfil de amplificación puede compararse con perfiles de amplificación de muestras que tienen cantidades conocidas de las bacterias a las que se dirigen los oligonucleótidos del conjunto de sondas o de cantidades conocidas de secuencia diana. En otras realizaciones el umbral de ciclo (C_T) puede usarse para calcular la cantidad de secuencia diana y por lo tanto las cantidades de la bacteria a la que se dirigen los oligonucleótidos del conjunto de sondas en la muestra. En todas las qPCR hay un umbral en el que la fluorescencia el producto de amplificación se detecta por encima del fondo. El ciclo en el que se cruza este umbral es el C_T . En la fase exponencial de la reacción la cantidad de ADN que teóricamente se duplica en cada ciclo y por lo tanto pueden calcularse cantidades relativas de ADN entre muestras comparando los valores de C_T que quedan en la fase exponencial. Si la comparación se realiza con muestras con una cantidad conocida de un molde, la cantidad de molde en la muestra de ensayo puede calcularse y la cantidad de secuencia diana de los oligonucleótidos del conjunto de sondas presente en la muestra y por lo tanto la cantidad de bacterias que contiene la secuencia diana de los oligonucleótidos del conjunto de sondas en la muestra puede determinarse.

Puede usarse en la práctica de la invención una combinación de una o más de las técnicas anteriormente descritas para determinar la cantidad de secuencias de nucleótidos diana.

El sujeto puede ser cualquier sujeto humano o animal no humano, pero más particularmente puede ser un vertebrado, por ejemplo un mamífero, incluyendo ganado y animales de compañía. Preferentemente el sujeto es un ser humano. El sujeto puede ser de cualquier edad, por ejemplo un lactante, un niño, un joven, un adolescente o un adulto, preferentemente un adulto. En seres humanos, se considera que un adulto tiene al menos 16 años de edad y un lactante tiene hasta 2 años de edad. En ciertas realizaciones el sujeto será un lactante, en otras será un niño o un adulto.

Los métodos de la invención son métodos *in vitro* realizados usando cualquier muestra tomada del tracto GI. El tracto GI, también denominado tracto digestivo o canal alimentario (y pudiendo usarse dichos términos indistintamente con tracto GI) es la serie continua de órganos que comienzan en la boca y terminan en el ano. Específicamente esta secuencia consiste en la boca, la faringe, el esófago, el estómago, el duodeno, el intestino delgado, el intestino grueso y el ano. Estos órganos pueden subdividirse en el tracto GI superior, que consiste en la boca, faringe, esófago, estómago y duodeno, y el tracto GI inferior, que consiste en el yeyuno, el íleon (juntos el intestino delgado), el ciego, el colon, el recto (juntos el intestino grueso) y el ano.

Una muestra de tracto GI de uso en la invención puede incluir, pero sin limitación cualquier fluido o sólido tomado del lumen o la superficie del tracto GI o cualquier muestra de cualquiera de los tejidos que forman los órganos del tracto GI. Por lo tanto la muestra puede ser cualquier contenido luminal del tracto GI (por ejemplo contenidos del estómago, contenidos del intestino, moco y heces, o combinaciones de los mismos) así como muestras obtenidas mecánicamente del tracto GI por ejemplo por hisopo, aclarado, aspirado o raspado de una cavidad o superficie del tracto GI o por biopsia de un tejido/órgano del tracto GI. Se prefieren muestras fecales. La muestra también puede obtenerse de parte de un tejido/órgano del tracto GI que se ha retirado quirúrgicamente. La muestra puede ser una parte del tejido/órgano escindido. En realizaciones en las que la muestra es una muestra de un tejido/órgano del tracto GI la muestra puede comprender una parte de la mucosa, la submucosa, la capa muscular externa, la adventicia y/o la serosa del tejido/órgano del tracto GI. Dichas muestras tisulares pueden obtenerse por biopsia durante un procedimiento endoscópico. Preferentemente la muestra se obtiene del tracto GI inferior, es decir del yeyuno, el íleon, el ciego, el colon, el recto o el ano. Más preferentemente la muestra es una muestra mucosa o luminal. Pueden recogerse muestras fecales por el hisopo, aclarado, aspirado o raspado del recto o el ano o, más sencillamente, la recogida de heces después de la defecación.

La muestra puede usarse en los métodos de la invención en la forma en que se recuperó inicialmente. La muestra también puede haber experimentado algún grado de manipulación, refinamiento o purificación antes de usarse en los métodos de la invención. Por lo tanto, el término "muestra" también incluye preparaciones de la misma, por ejemplo materiales de partida relativamente puros o parcialmente purificados, tales como preparaciones semipuras de las muestras anteriormente mencionadas. El término "muestra" también incluye preparaciones de las muestras anteriormente mencionadas en las que su ARN, incluyendo el ARNr de 16S, ha experimentado transcripción inversa.

La purificación puede ser ligera, por ejemplo sumando no más que la concentración de los sólidos, o células, de la muestra en un volumen menor o la separación de células de parte de o todo el resto de la muestra. Se describen técnicas de aislamiento celular representativas en los documentos WO98/51693 y WO01/53525.

En otras realizaciones la invención usa una preparación del ácido nucleico de las muestras anteriormente mencionadas, preferentemente una preparación en la que se han marcado los ácidos nucleicos. Dichas preparaciones incluyen productos de transcripción inversa y/o productos de amplificación de dichas muestras o preparaciones de ácido nucleico de las mismas. Preferentemente el ácido nucleico predominante de la preparación de ácido nucleico es ADN.

Las técnicas para el aislamiento de ácido nucleico de muestras, incluyendo muestras complejas, son numerosas y se conocen bien en la técnica y se describen con detalle en la bibliografía. Las técnicas descritas en los documentos WO98/51693 y WO01/53525 también pueden emplearse para preparar ácidos nucleicos de las muestras anteriormente mencionadas. Estas preparaciones incluyen preparaciones de ácido nucleico relativamente puras o parcialmente purificadas.

Preferentemente la reacción de amplificación realizada en la muestra será universal, o sustancialmente universal, porque el ácido nucleico para amplificar, es decir la región de ARNr de 16S o ARNr de 16S que incorpora las secuencias diana anteriormente realizadas, se amplifica de todas, o al menos sustancialmente todas, las células procariotas que podrían estar presentes en una muestra. La expresión "amplificación de sustancialmente todas las células procariotas presentes en una muestra" se refiere al número de especies diferentes de células procariotas en la muestra que tendrán el ácido nucleico para amplificar, amplificado. Por lo tanto, en esta realización el ácido nucleico para amplificar se amplifica a partir de al menos una especie representativa de sustancialmente todas las especies de células procariotas en la muestra.

Por "célula procariota" se entiende cualquier organismo que carezca de un núcleo celular, es decir cualquier organismo de los dominios *Bacteria* y *Archaea*.

Convenientemente esta amplificación universal puede realizarse usando un cebador directo que se dirige a la región conservada entre V2 y V3 (por ejemplo, la descrita en Nadkarni *et al.*, 2002. Microbiology 148, 257-266) con un cebador inverso que se dirige al extremo 3' del gen de ARNr de 16S (por ejemplo, el descrito en Weisburg *et al.*, 1991, J Bacteriol 173, 697-703). En otras realizaciones esta amplificación universal puede realizarse usando un par de cebadores que tienen las secuencias TCC TAC GGG AGG CAG (SEQ ID NO 53), también denominado MangalaF-1, y CGG TTA CCT TGT TAC GAC TT (SEQ ID NO 54), también denominado 16SU1510R. Este par de cebadores se describe en más detalle en el documento US 2011/0104692.

La secuencia de nucleótidos diana para amplificar en esta realización está presente por lo tanto en ARNr de 16S y el gen de ARNr 16S correspondiente (ADNr). Por lo tanto, la referencia a la amplificación de esta secuencia de nucleótidos diana es una referencia a un aumento del número de ácidos nucleicos que contienen esa secuencia de nucleótidos sin limitación en el tipo de ácidos nucleicos que contiene la secuencia de nucleótidos. Preferentemente estos ácidos nucleicos estarán marcados. Típicamente, el ácido nucleico que se forma como el producto de amplificación es ADN, aunque la secuencia de nucleótidos contenida en ese ácido nucleico aún será la misma que la de la secuencia de nucleótidos diana, o el complemento de la misma.

Convenientemente, esta realización de la invención se realizará con ADNr 16S, por ejemplo un gen de ARNr 16S, como el molde.

En otras realizaciones ARNr 16S puede ser la fuente de la secuencia de nucleótidos diana para amplificar. Cuando una secuencia de nucleótidos diana de ARNr 16S se amplifica en esta realización del método de la invención habrá una etapa en la que una ADN polimerasa dependiente de ARN cataliza la formación de una molécula de ADN complementaria del molde de ARNr 16S (ADNc). Este proceso se denomina "transcripción inversa". Más específicamente la ADN polimerasa dependiente de ARN cataliza la polimerización de desoxirribonucleósido trifosfatos en una secuencia que es complementaria (es decir siguiendo las reglas de formación de pares de bases de Watson-Crick) de una secuencia de ARNr de molde con cebador.

Se han identificado numerosas enzimas que tienen la capacidad de catalizar esta reacción y los ejemplos incluyen, pero sin limitación transcriptasa inversa de VIH, transcriptasa inversa de AMV, transcriptasa inversa de M-MLV, polimerasa de *C. therm.* y polimerasa Tth. En su nivel más básico una mezcla de reacción de transcripción inversa completa contendrá una enzima de transcripción inversa, el molde de ARNr, cebadores adecuados que pueden unirse con el molde y a partir de los que la transcriptasa inversa puede comenzar la polimerización, dNTP y un tampón adecuado. La incubación de la mezcla a la temperatura de trabajo de la transcriptasa inversa da como resultado producción de ADNc.

Tras completar la reacción de transcripción inversa el ADNc puede usarse como el molde en la realización del método de la invención descrito anteriormente. El ADNc por lo tanto tiene una secuencia de nucleótidos que es complementaria de la molécula de ARNr que era su molde. Además el ADNc tiene una secuencia de nucleótidos que es igual que una secuencia de nucleótidos contenida en una cadena del gen del molde de ARNr y el ADNc es complementario con una secuencia de nucleótidos contenida en la otra cadena del gen del molde de ARNr.

Como se ha mencionado anteriormente, en realizaciones del método de la invención en las que se amplifica ácido nucleico en una etapa precedente, si se usa ARNr 16S como la fuente de la secuencia de nucleótidos diana (a diferencia de ADNr 16S, por ejemplo un gen de ARNr 16S) se requiere una etapa de transcripción inversa inicial. Las reacciones de amplificación ligadas a transcripción inversa, en particular PCR, pueden ser procesos de "una etapa" o "dos etapas". En un proceso de una etapa los componentes de la reacción de transcripción inversa y la reacción de amplificación de ácido nucleico están presentes en un único recipiente de reacción y típicamente las condiciones de reacción temprana se seleccionan para permitir que la reacción de transcripción inversa continúe hasta completarse y las condiciones de reacción se cambian después a condiciones adecuadas para permitir que continúe la reacción de amplificación de ácido nucleico.

En un proceso de dos etapas los componentes de la reacción de transcripción inversa se combinan en primer lugar y se realiza la reacción de transcripción inversa. El producto de transcripción inversa se combina después con los componentes de la reacción de amplificación y se somete a la reacción de amplificación. En un protocolo de dos etapas "en un tubo" los componentes de la reacción de amplificación se añaden al mismo recipiente de reacción en el que se realizó la reacción de transcripción inversa. En un protocolo de dos etapas en "dos tubos" la reacción de amplificación se realiza en un recipiente de reacción nuevo.

Se cree que muchas enfermedades y afecciones, o estadios de las mismas, están ligadas a perfiles característicos de la microbiota del tracto GI o la microbiota de las regiones/partes del mismo, por ejemplo los descritos anteriormente. En algunos casos la enfermedad o afección puede estar provocada por, o empeora por, perturbaciones en el perfil de la microbiota del tracto GI o de regiones/partes del mismo. En otros casos la enfermedad o afección provoca, o por algún mecanismo da como resultado, la presentación de un perfil particular de la microbiota del tracto GI o regiones/partes del mismo. En consecuencia, analizando perfiles de microbiota en muestras del tracto GI, puede proporcionarse información que permite el diagnóstico de una enfermedad o afección, o que permite una evaluación del riesgo de desarrollar una enfermedad o afección, que se ha determinado que se caracteriza por un perfil de microbiota particular. El conjunto de sondas de la invención puede usarse por lo tanto para preparar perfiles de microbiota convencionales del tracto GI que son característicos de una enfermedad o afección, o estadio de la misma, o el riesgo de desarrollar una enfermedad o afección. El perfil también puede usarse en pronóstico de enfermedad y la supervisión de enfermedad.

Por lo tanto en otro aspecto la invención proporciona un método para preparar un perfil de microbiota convencional del tracto GI que es característico de una enfermedad o afección o un estadio de la misma o el riesgo de desarrollar una enfermedad o afección, comprendiendo dicho método:

(i) identificar un sujeto con dicha enfermedad o afección o estadio de la misma o que está en riesgo de desarrollar dicha enfermedad o afección y poner en contacto una muestra del tracto GI de dicho sujeto con un conjunto de sondas oligonucleotídicas como se define en las reivindicaciones 1-7;

(ii) someter la muestra y el conjunto de sondas a condiciones que permiten la hibridación de las sondas con sus secuencias de nucleótidos diana dentro de moléculas de ácido nucleico en dicha muestra; y

(iii) para cada oligonucleótido en dicho conjunto de sondas, determinar la cantidad de su secuencia de nucleótidos diana que está presente en dicha muestra.

Un perfil de la microbiota del tracto GI del sujeto se genera por lo tanto a partir de las cantidades de la secuencia diana para cada oligonucleótido presente en dicha muestra y el perfil es característico de dicha enfermedad o afección (o estadio de la misma) o el riesgo de desarrollar dicha enfermedad o afección.

Una vez que se ha obtenido un perfil convencional para una enfermedad o afección particular o riesgo de desarrollarla, típicamente después de la realización del perfil de una pluralidad de muestras de una pluralidad de sujetos con la misma enfermedad o afección o estadio de la misma, este perfil puede usarse como la base de un proceso de diagnóstico para determinar si un sujeto adicional padece, o está en riesgo de desarrollar, dicha enfermedad o afección, o para determinar el progreso o alcance de la aparición de dicha enfermedad o afección para fines de pronóstico. El perfil convencional puede proporcionarse digitalmente, por ejemplo en medio digital o mediante transferencia electrónica al usuario. En otras realizaciones puede establecerse un sistema en el que el perfil obtenido del sujeto que se ensaya contribuye al desarrollo del perfil convencional.

En un aspecto adicional la invención proporciona un método para diagnosticar o supervisar una enfermedad o afección en un sujeto o predecir o evaluar el riesgo de un sujeto de desarrollar una enfermedad o afección, comprendiendo dicho método:

(a) realizar el perfil de la microbiota del tracto GI de un sujeto con un método como se define en una cualquiera de las reivindicaciones 8 a 17

(b1) comparar dicho perfil con un perfil de microbiota convencional del tracto GI que es característico de una enfermedad o afección o un estadio de la misma o el riesgo de desarrollar una enfermedad o afección y/o

(b2) comparar dicho perfil con un perfil de microbiota anterior del tracto GI del sujeto; y

(c) determinar el grado de correlación entre dichos perfiles.

En esta realización dicho grado de correlación es indicativo de la presencia o ausencia de dicha enfermedad o afección, o el riesgo de desarrollar dicha enfermedad o afección, o el progreso de la enfermedad o afección.

En un aspecto adicional la invención también proporciona un método para diagnosticar o supervisar una enfermedad o afección en un sujeto o predecir o evaluar el riesgo de un sujeto de desarrollar una enfermedad o afección, comprendiendo dicho método:

(a) comparar los resultados de un método como se describe en una cualquiera de las reivindicaciones 8 a 17 con un perfil de microbiota convencional del tracto GI que es característico de una enfermedad o afección o un estadio de la misma o el riesgo de desarrollar una enfermedad o afección y/o

(a2) comparar los resultados de un método como se ha descrito anteriormente con un perfil de microbiota anterior del tracto GI del sujeto, y

(b) determinar el grado de correlación entre dichos perfiles,

en el que el grado de correlación es indicativo de la presencia o ausencia de dicha enfermedad o afección, o el riesgo de desarrollar dicha enfermedad o afección, o el progreso de la enfermedad o afección.

Preferentemente el perfil con el que se compara el perfil del sujeto que se ensaya será un perfil preparado de acuerdo con la invención. Este puede ser un perfil preparado previamente, o podría ser un perfil preparado al mismo tiempo o sustancialmente al mismo tiempo que se analiza la muestra que se investiga.

"Diagnóstico" se refiere a la determinación de la presencia o existencia de una enfermedad o afección o estadio de la misma en un organismo. "Supervisión" se refiere al establecimiento del alcance de, o posibles cambios en, una enfermedad o afección, particularmente cuando se sabe que un individuo padece una enfermedad o afección, por ejemplo supervisar los efectos del tratamiento o el desarrollo de una enfermedad o afección, por ejemplo determinar la idoneidad de un tratamiento, proporcionar un pronóstico y/o determinar si un paciente está en remisión o recaída.

"Evaluar el riesgo de un sujeto de desarrollar una enfermedad o afección" se refiere a la determinación de la probabilidad o verosimilitud de que el sujeto desarrolle la enfermedad o afección. Esto puede expresarse como una probabilidad numérica en algunas realizaciones. La evaluación del riesgo puede ser en virtud del grado en que se ve una correlación entre el perfil de una muestra de un sujeto que se investiga y el perfil de una enfermedad o afección, o la correlación entre el perfil de una muestra de un sujeto que se investiga y el perfil que se considera característico de una muestra de un sujeto determinado como con un nivel particular de riesgo de desarrollar una enfermedad o afección.

"Enfermedad" se refiere a un estado de alteración patológica en relación con lo normal que puede resultar, por ejemplo, de infección o una imperfección genética adquirida o congénita.

Una "afección" se refiere a un estado de la mente o el cuerpo de un organismo que no ha sucedido mediante enfermedad, por ejemplo la presencia de un agente en el cuerpo tal como una toxina, fármaco o contaminante, o embarazo.

"Estadio del mismo" se refiere a diferentes estadios de una enfermedad o afección que pueden mostrar o no cambios metabólicos o fisiológicos particulares, pero no muestran cambios en el perfil de la microbiota del tracto GI. En algunas realizaciones las diferencias observadas en el perfil de la microbiota del tracto GI pueden conducir a clasificación previamente inapreciada del progreso de una enfermedad o afección.

Los datos generados usando los métodos anteriormente mencionados pueden analizarse usando diversas técnicas, de la representación visual más básica (por ejemplo en relación con la intensidad de señal) a manipulación de datos más compleja, que puede cuantificarse y expresarse matemáticamente, para preparar los perfiles de microbiota del tracto GI que reflejan la interrelación de los niveles relativos de cada secuencia diana con la que se unen las diversas sondas (y por lo tanto los niveles relativos de bacterias que contiene la secuencia diana con la que se unen las diversas sondas). Convenientemente, los datos sin procesar generados de este modo pueden manipularse por procesamiento de datos y métodos estadísticos, particularmente normalización y estandarización de los datos, y consultando los datos estadísticamente para determinar si dichos datos reflejan el perfil de una enfermedad particular, acción o estadio de la misma. El experto en la materia sería consciente de técnicas estadísticas adecuadas para usar. Preferentemente la técnica estadística proporcionará un "P valor" como un indicio de que la tendencia que se observa no es una tendencia aleatoria. Un resultado estadísticamente significativo, es decir un resultado que no es atribuible a variación aleatoria en comparación con su control, tendrá un P valor de $<0,05$, preferentemente $<0,01$, $<0,005$ o $<0,001$. Únicamente como ejemplo, son técnicas adecuadas para medir la significación estadística en los métodos de la invención ANOVA, ensayo de Mann-Whitney-Wilcoxon (MWW), ensayo de Kruskal-Wallis y ensayo de diferencias honestamente significativas de Tukey (HSD). El experto en la materia estará familiarizado con muchas otras. En algunas realizaciones un ensayo de permutación podría ser apropiado, por ejemplo el descrito en Langsrud (2002, Journal Of The Royal Statistical Society Series D 51, 305-317).

Las enfermedades y afecciones que pueden investigarse usando los métodos de la invención no están limitadas, aunque los aspectos de diagnóstico de la invención se basan en que exista la presencia de un perfil uniforme de microbiota de tracto GI que es característico de la enfermedad o afección que se investiga. Es muy probable que las enfermedades y afecciones que afectan al tracto GI den como resultado perfiles de microbiota característicos, por ejemplo enfermedad inflamatoria del intestino (EII), enfermedad de Crohn (EC), colitis ulcerosa (CU), síndrome del intestino irritable (SII) y cánceres del tracto GI (por ejemplo cáncer de la boca, faringe, esófago, estómago, duodeno, yeyuno, íleon, ciego, colon, recto o ano) y también existen pruebas de relaciones entre la microbiota del tracto GI y enfermedades y afecciones que se considera que no están relacionadas con el tracto GI, por ejemplo las enfermedades atópicas, por ejemplo, eccema, asma, dermatitis atópica, conjuntivitis alérgica, rinitis alérgica y alergias alimentarias; trastornos metabólicos, por ejemplo diabetes mellitus (tipo 1 y tipo 2), obesidad y síndrome metabólico; trastornos neurológicos, por ejemplo depresión, esclerosis múltiple, demencia y enfermedad de Alzheimer; y autismo. Cualquiera de estas enfermedades o afecciones puede diagnosticarse o supervisarse de acuerdo con la invención.

El método de diagnóstico puede usarse solo como una alternativa a otras técnicas de diagnóstico o además de dichas técnicas. Por ejemplo, pueden usarse métodos de la invención como una medida de diagnóstico alternativa o adicional al diagnóstico usando técnicas de captura de imágenes tales como imagen por resonancia magnética (IRM), captura de imágenes por ultrasonido, captura de imágenes nuclear, captura de imágenes por rayos X o endoscopia.

Por lo tanto, en un aspecto adicional, la presente invención proporciona un método para obtener información relevante para el diagnóstico o la supervisión de una enfermedad o afección o la evaluación del riesgo de desarrollar una enfermedad o afección que comprende un método de realizar el perfil de la microbiota del tracto GI de un sujeto como se define en las reivindicaciones 8-15.

En un aspecto adicional la invención proporciona kits que comprenden el conjunto de sondas como se define en las reivindicaciones 1-7.

En un aspecto adicional la invención proporciona el uso del conjunto de sondas definido en las reivindicaciones 1-7 en la fabricación de los kits.

Los kits de la invención pueden diseñarse para su uso en los métodos de la invención y pueden comprender componentes adicionales. Cada componente puede proporcionarse en un compartimento o recipiente separado. Cuando sea conveniente y práctico, podrían proporcionarse mezclas de componentes. Los componentes pueden proporcionarse en forma seca, por ejemplo cristalizada, criodesecada o liofilizada, o en solución, típicamente dichas composiciones líquidas serán acuosas y se tamponarán con un tampón convencional tal como Tris, HEPES, etc.

El kit también puede proporcionarse con instrucciones para usar el kit, o con directrices sobre cómo obtener instrucciones.

Los componentes adicionales pueden ser cualquiera de los diversos componentes que pueden usarse para llevar a cabo los métodos de la invención, por ejemplo cualquier componente analizado anteriormente. En una realización preferida el kit comprende además medios para marcaje selectivo de las sondas oligonucleotídicas. En una realización preferida el kit comprende además soportes sólidos adecuados en los que pueden inmovilizarse los oligonucleótidos del conjunto de sondas de la invención, por ejemplo cualquiera de los soportes sólidos descritos en el presente documento. En otras realizaciones algunos o todos los oligonucleótidos del conjunto de sondas de la invención se proporcionan en el kit en forma inmovilizada.

Componentes adicionales podrían ser opcionalmente cualquiera o todos de los medios, por ejemplo tampones, enzimas, etc. para realizar una amplificación y/o reacción de extensión de cebadores con los oligonucleótidos de la invención. Por ejemplo, los kits pueden contener opcionalmente un tampón de reacción de PCR, nucleótido trifosfatos, cebadores oligonucleotídicos adicionales o ADN polimerasas, preferentemente una polimerasa termoestable tal como Taq polimerasa.

Componentes adicionales podrían ser opcionalmente ser cualquiera de o todos los medios, por ejemplo tampones, enzimas, etc. para realizar una reacción de transcripción inversa. Por ejemplo una transcriptasa inversa, cebadores específicos de ARN, un tampón de reacción RT y nucleótido trifosfatos.

Componentes adicionales podrían ser opcionalmente cualquiera de o todos los medios para tomar la muestra. Por ejemplo dichos medios podrían incluir varillas, aparato de biopsia, dispositivos de hisopado, bolsas o recipientes. Preferentemente estos medios se proporcionarían en forma estéril.

Componentes adicionales podrían ser opcionalmente cualquiera de o todos los medios para purificar o refinar la muestra. Por ejemplo, medios para aislar o concentrar células en una muestra, por ejemplo soportes sólidos de unión a células o dispositivos de filtración. En otras realizaciones los medios para purificar o refinar la muestra podrían ser cualquiera de o todos los medios para extraer ácido nucleico de una muestra. Por ejemplo reactivos de lisis celular (por ejemplo sales caotrópicas, alcoholes, detergentes, compuestos de alteración de membrana), soportes sólidos de unión a ácido nucleico (por ejemplo como se ha descrito anteriormente) o agentes de precipitación de ácido nucleico (por ejemplo sales, alcoholes).

Componentes adicionales podrían ser opcionalmente ser cualquiera de o todos los medios para detectar ácido nucleico amplificado. Por ejemplo los marcadores descritos en el presente documento (por ejemplo colorantes de unión a ADN bicatenario, sondas oligonucleotídicas marcadas), aparatos para detectar estos marcadores, materiales y aparatos de electroforesis, o materiales y aparatos de cromatografía.

Componentes adicionales podrían ser además oligonucleótidos que hibridan selectivamente con ácidos nucleicos diana indicativos de cualquier otra enfermedad o afección médica, particularmente afecciones asociadas con la microbiota gastrointestinal (por ejemplo EC, SII, CU, EII) y que pueden usarse en consecuencia de una manera similar a los oligonucleótidos de la invención para proporcionar información relevante para un diagnóstico de cualquier otra enfermedad de condición médica, particularmente afecciones asociadas con la microbiota gastrointestinal (por ejemplo EC, SII, CU, EII). Estos oligonucleótidos pueden considerarse parte del conjunto de sondas de la invención.

Se describen los siguientes ejemplos no limitantes en los que:

La Figura 1 muestra el desarrollo temporal de filos bacterianos en lactantes sensibilizados y no sensibilizados. La señal promedio logarítmica para cada sonda se muestra como línea continua, mientras que la señal logarítmica de todos los puntos temporales medidos se muestra como puntos (niveles por encima de un umbral de señal de 50, indicado por líneas punteadas). Las líneas gris oscuro y los puntos son para niños sensibilizados, mientras que las líneas gris claro y los puntos son para niños no sensibilizados. Los valores < 0 se ajustan a 0,001 antes de la transformación logarítmica.

La Figura 2 muestra el desarrollo temporal de géneros/especies bacterianos en lactantes sensibilizados y no sensibilizados. La señal promedio logarítmica para cada sonda se muestra como línea continua, mientras que la señal logarítmica de todos los puntos temporales medidos se muestra como puntos (niveles por encima de un umbral de señal de 50 se indican por líneas punteadas). Las líneas gris oscuro y los puntos son para niños sensibilizados, mientras que las líneas gris claro y los puntos son para niños no sensibilizados. Los valores < 0 se ajustan a 0,001 antes de la transformación logarítmica. No se muestran las sondas con máxima señal por debajo del umbral.

Tabla 1

Secuencia de sonda	SEQ ID NO	Bacterias diana (descripción general)
TTGCGGCTCAACCGTAAAATTG	SEQ ID NO 5	<i>Bacteroides</i>
CAATTTTACGGTTGAGCCGCAA	SEQ ID NO 31	<i>Bacteroides</i>
GCACTCAAGACATCCAGTATCAACTG	SEQ ID NO 6	<i>Bacteroides (dorei, fragilis, thetaiotaomicron, vulgatus)</i>
CAGTTGATACTGGATGTCTTGAGTGC	SEQ ID NO 32	<i>Bacteroides (dorei, fragilis, thetaiotaomicron, vulgatus)</i>
AGGGCAGTCATCCTTCACG	SEQ ID NO 7	<i>Bacteroides (dorei, fragilis, thetaiotaomicron, vulgatus)</i>
CGTGAAGGATGACTGCCCT	SEQ ID NO 33	<i>Bacteroides (dorei, fragilis, thetaiotaomicron, vulgatus)</i>
CAATCGGAGTTCTTCGTGATATCTAAG	SEQ ID NO 8	<i>Bacteroides</i>
CTTAGATATCACGAAGAACTCCGATTG	SEQ ID NO 34	<i>Bacteroides</i>

Tabla 2

Secuencia de sonda	SEQ ID NO	Bacterias diana (descripción general)
TGTTGTGGTTAATAACCGCAGCAATTGA	SEQ ID NO 9	<i>Salmonella, Enterobacter, Citrobacter, Cronobacter</i>
TCAATTGCTGCGGTTATTAACCACAACA	SEQ ID NO 35	<i>Salmonella, Enterobacter, Citrobacter, Cronobacter</i>
TCCAATGACCCTCCC	SEQ ID NO 10	<i>Enterococcus, Listeria</i>
GGGAGGGTCATTGGA	SEQ ID NO 36	<i>Enterococcus, Listeria</i>
CACTCTCACACCCGTT	SEQ ID NO 11	<i>Streptococcus sanguinis</i>
AACGGGTGTGAGAGTG	SEQ ID NO 37	<i>Streptococcus sanguinis</i>
GTTGCTCGGTCAGACTT	SEQ ID NO 12	<i>Streptococcus, Enterococcus</i>
AAGTCTGACCGAGCAAC	SEQ ID NO 38	<i>Streptococcus, Enterococcus</i>
CGTGGCTTTCTGATTAGGTA	SEQ ID NO 13	<i>Staphylococcus</i>
TACCTAATCAGAAAGCCACG	SEQ ID NO 39	<i>Staphylococcus</i>
TGCTTATTCAACGGGTAACT	SEQ ID NO 14	<i>Bifidobacterium longum</i>
AGTTTACCCGTTGAATAAGCA	SEQ ID NO 40	<i>Bifidobacterium longum</i>
CCGTCACTCGGCTACCATTC	SEQ ID NO 15	<i>Clostridium ramosum</i>
GAAATGGTAGCCGAGTGACGG	SEQ ID NO 41	<i>Clostridium ramosum</i>
GATTTTCCACTCCCACCAT	SEQ ID NO 16	<i>Streptococcus pyogenes</i>
ATGGTGGGAGTGAAAATC	SEQ ID NO 42	<i>Streptococcus pyogenes</i>
CCGTCAAGGGACAAG	SEQ ID NO 17	<i>Listeria monocytogenes</i>
CTTGTCCTTGACGG	SEQ ID NO 43	<i>Listeria monocytogenes</i>
CGGTGCTTATTCGAAAGGTACACT	SEQ ID NO 18	<i>Bifidobacterium breve</i>
AGTGTACCTTTCAATAAGCACCG	SEQ ID NO 44	<i>Bifidobacterium breve</i>

Tabla 3

Secuencia de sonda	SEQ ID NO	Bacteria diana (descripción general)
CGCCTGCCTCAAACATA	SEQ ID NO 19	<i>Parabacteroides</i>
TATGTTTGAGGCAGGCG	SEQ ID NO 45	<i>Parabacteroides</i>
CAGGTGTAGCGGTGAAATGCGTAGAGAT	SEQ ID NO 20	Gamma-proteobacteria
ATCTCTACGCATTTCACCGCTACACCTG	SEQ ID NO 46	Gamma-proteobacteria
ACGCTCGCACC	SEQ ID NO 21	<i>Haemophilus</i>
GGTGCGAGCGT	SEQ ID NO 47	<i>Haemophilus</i>
CGGGGATTTACATCTGA	SEQ ID NO 22	Subgrupo de Gamma-proteobacteria
TCAGATGTGAAATCCCCG	SEQ ID NO 48	Subgrupo de Gamma-proteobacteria
TGCCAGTTTCGAATGCAGTT	SEQ ID NO 23	Subgrupo de Gamma-proteobacteria
AACTGCATTGAACTGGCA	SEQ ID NO 49	Subgrupo de Gamma-proteobacteria
GTGCTTCTTCTGCGGGTAA	SEQ ID NO 24	Subgrupo de Gamma-proteobacteria
TTACCCGCAGAAGAAGCAC	SEQ ID NO 50	Subgrupo de Gamma-proteobacteria
GCTACACATGGAGTTCCA	SEQ ID NO 25	Subgrupo de <i>Lactobacillus</i>
TGGAAGTCCATGTGTAGC	SEQ ID NO 51	Subgrupo de <i>Lactobacillus</i>
CGTAGTTAGCCGTGG	SEQ ID NO 26	<i>Clostridium neonatale</i>
CCACGGCTAACTACG	SEQ ID NO 52	<i>Clostridium neonatale</i>

Ejemplo 1

- 5 El objetivo del presente trabajo fue comparar de forma prospectiva el desarrollo de la microbiota dominante en niños sensibilizados y niños no sensibilizados a IgE durante los primeros dos años de vida. Para conseguir esto y otras tareas relacionadas con la microbiota del intestino, era necesaria una herramienta para explorar rápidamente la complejidad y composición de la bacteria en muestras de heces. Se desarrolló por lo tanto una micromatriz de gen de ARNr 16S de alto rendimiento de lactante, denominado ensayo de lactantes de GA-map. Los análisis de
- 10 micromatrices se realizaron en un subconjunto seleccionado de la cohorte de IM-PACT. Se seleccionó IgE específico como un marcador de atopía, ya que se mostró previamente que este marcador está correlacionado con las bacterias del intestino (resultados no publicados).

- 15 La diferencia principal entre la matriz de lactantes de GA-map y enfoques de matriz de gen de ARNr 16S alternativos es el uso de sondas de extensión de cebadores de un único nucleótido (SNUPE) altamente específicas para diferenciación de diana/no diana. La alta especificidad del ensayo de SNUPE se obtiene por incorporación basada en ADN polimerasa de un didesoxinucleótido marcado con fluorescencia. Las sondas de SNUPE se construyen de modo que las sondas hibriden adyacentes a posiciones génicas diferenciadoras. Para reducir la complejidad y para aumentar el rendimiento, el ensayo de lactantes de GA-map se dirigió a bacterias que se esperaba que colonizaran
- 20 el intestino del lactante. Las sondas se seleccionaron basándose en el criterio del número mínimo de sondas que abarcan la diversidad esperada de bacterias en el intestino del lactante.

Materiales y métodos25 *Cohorte*

- El Estudio de Prevención de la Alergia Entre Niños en Trondheim (PACT) es un estudio de intervención basado en población grande en Noruega centrado en la alergia infantil. Las muestras incluidas aquí son un subconjunto del estudio de PACT, en el que se tomaron medidas de inmunología y microbiología. Para la familia de subestudio
- 30 participaron doctores y matronas en Trondheim en el reclutamiento de una población no seleccionada de mujeres durante comprobaciones de embarazo temprano ordinarias hasta que se habían aprobado 720 para su participación. Las mujeres rellenaron cuestionarios sobre factores de riesgo durante el embarazo, a las seis semanas después del parto, al año y a los dos años después de dar a luz. Las preguntas fueron sobre alergia en la familia, condiciones de alojamiento, dieta y estilo de vida, y después del nacimiento sobre lactancia, complementos alimentarios, dieta, infecciones, vacunas, antibióticos, estancias en guarderías y exposición a nicotina. Cuando los lactantes cumplieron
- 35 dos años, se presentó otro cuestionario sobre salud y enfermedad. Se evaluó la sensibilización atópica como IgE específico elevado ($> 0,35$ kU/ml) en suero usando un ensayo para una serie de alérgenos (sistema de IgE

específico de alérgenos Immulite 2000, Siemens Medical Solutions Diagnostics). La cohorte se analizó inicialmente con respecto a doce bacterias específicas por PCR cuantitativa (resultados no publicados). Aquí, se seleccionó una serie de lactantes para ensayos de matriz de lactante de GA-map en profundidad en varias muestras, y estado de sensibilización. Se seleccionó un total de 16 niños sensibilizados y 31 no sensibilizados, que representan un total de 216 muestras fecales.

Preparación de muestras y amplificación por PCR

Se recogieron heces del pañal y se transfirieron a un medio de transporte Carry Blair por los padres, se almacenó inmediatamente a -18 °C en casa antes de transportarse a almacenamiento permanente a -80 °C hasta su análisis posterior. Se usó lisis mecánica para rotura celular, y se usó un método basado en perlas magnéticas automáticas para purificación de ADN.

Se combinó el uso de un cebador directo que se dirigía a la región conservada entre V2 y V3 (Nadkarni *et al.*, 2002. Microbiology 148, 257-266) con un cebador inverso que se dirigía al extremo 3' del gen de ARNr 16S (Weisburg *et al.*, 1991, J Bacteriol 173, 697-703). Se usaron 1,5 U de HotFirePol (Solis Biodyne, Tartu, Estonia), tampón B2 1 x (Solis Biodyne), MgCl₂ 2,5 mM (Solis Biodyne), dNTP 200 µM (Thermo Fisher Scientific, Waltham, Estados Unidos), 0,2 µM de cada cebador directo e inverso y aproximadamente 10 a 50 ng de molde en un volumen total de 25 µl. Se amplificó una de las muestras tres veces para examinar la reproducibilidad de la reacción de PCR (descrita en más detalle en la sección de electroforesis capilar). El protocolo de la amplificación incluyó un estadio de activación de 15 min a 95 °C, seguido de 30 ciclos con 30 s de desnaturalización a 95 °C, 30 s de hibridación a 55 °C y 90 s de extensión a 72 °C. Se incluyó una elongación final durante 7 min a 72 °C para completar todos los productos de PCR.

Para los ensayos iniciales de la matriz, se realizó PRC de gen de ARNr 16S en ADN bacteriano a partir de cultivos puros de 26 cepas enumeradas en la Tabla 4, y los productos de PCR se ensayaron en el ensayo de lactantes de GA-map corriente abajo. Las cepas se secuenciaron para confirmar su identidad y posibles mutaciones (se enumeran números de referencia de secuencia en la Tabla 4). Se incluyó un control positivo que consistía en una mezcla de ADN de cultivos puros de 8 cepas bacterianas relevantes así como un control negativo que consiste en H₂O durante la reacción de PCR de gen de ARNr 16S y el ensayo de lactantes de GA-map corriente abajo. Los controles positivos se usaron como un control de calidad de la reacción de marcaje e hibridación de las matrices (resultados no mostrados)

Diseño del ensayo de lactante de GA-map

El ensayo de GA-map se basa en el principio de extensión de un único nucleótido (SNupE) en combinación con hibridación de micromatriz (Rudi *et al.*, 1998, Appl Environ Microbiol 64, 2639-2643).

Las cepas bacterianas mostradas en la Tabla 4 se usaron para validación de sondas. Para la construcción de sondas se usó un conjunto de datos combinado que consistía en un total de 3580 secuencias génicas de ARNr 16S (Palmer *et al.*, 2007, PLoS Biology 5, e177; Rudi *et al.*, 2007, Appl Environ Microbiol 73, 2727-2734), además de un conjunto de patógenos conocidos.

Tabla 4. Cepas bacterianas usadas para evaluación de sondas

Clase	Especies	Cepa	N.º de referencia
Actinobacteria	<i>Bifidobacterium breve</i>	DSM20213	HQ012023
	<i>Bifidobacterium longum subsp. infantis</i>	DSM20088	HQ012021
	<i>Bifidobacterium longum subsp. longum</i>	DSM20219	HQ012022
Bacteroidetes	<i>Bacteroides dorei</i>	DSM17855	HQ012025
	<i>Bacteroides fragilis</i>	DSM2151	HQ012027
	<i>Bacteroides thetaiotaomicron</i>	DSM2079	HQ012026
	<i>Bacteroides vulgatus</i>	DSM1447	HQ012024
	<i>Parabacteroides distasonis</i>	DSM 20701	N/D
Firmicutes	<i>Clostridium periringens</i>	DSM756	HQ012013
	<i>Clostridium ramosum</i>	DSM1402	HQ012012
	<i>Enterococcus faecalis</i>	DSM20478	HQ012029
	<i>Enterococcus faecium</i>	DSM20477	HQ012007

Clase	Especies	Cepa	N.º de referencia
	<i>Lactobacillus acidophilus</i>	DSM20079	HQ012028
	<i>Lactobacillus rhamnosus</i>	DSM20021	HQ012008
	<i>Listeria monocytogenes</i>	DSM20600	HQ012006
	<i>Staphylococcus aureus subsp. aureus</i>	DSM20231	HQ012011
	<i>Streptococcus pneumoniae</i>	DSM20566	HQ012009
	<i>Streptococcus pyogenes</i>	DSM20565	HQ012030
	<i>Streptococcus sanguinis</i>	DSM20567	HQ012010
	<i>Veillonella atípica</i>	DSM20739	HQ012015
	<i>Veillonella dispar</i>	DSM20735	HQ012014
Proteobacteria	<i>Escherichia coli</i>	DSM30083	HQ012019
	<i>Haemophilus parainfluenzae</i>	DSM8978	HQ012020
	<i>Klebsiella pneumoniae subsp. pneumoniae</i>	DSM30104	HQ012018
	<i>Salmonella bongori</i>	DSM13772	HQ012016
	<i>Salmonella enterica subsp. enterica</i>	DSM17058	HQ012017

Se usó un proceso de cuatro etapas en el diseño de las sondas. 1) En primer lugar, se definió un conjunto de grupos diana y no diana basándose en un sistema de clasificación de coordenadas. 2) La segunda etapa fue identificar sondas que satisficaran los criterios de detección de diana y exclusión no de diana. Esto se basó en un criterio combinado de hibridación y marcaje. Todas las sondas se diseñaron con Tm mínima de 60 °C para el grupo diana, mientras que el no diana debería tener una Tm de < 30 °C, o la ausencia de una citosina como el nucleótido adyacente al extremo 3' de la sonda. Se identificaron todas las sondas que satisficaran los criterios. 3) Después se evaluaron las sondas de automarcaje o marcaje cruzado potenciales, además del potencial para hibridación cruzada en la matriz. 4) Finalmente, combinando el conocimiento acerca de grupos diana/no diana y la compatibilidad para cada una de las sondas se diseñaron matrices finales usando un enfoque jerárquico.

Se incluyó una sonda génica de ARNr 16S universal (UNI01) en los conjuntos de sondas para medir la abundancia total de ADN bacteriano en la muestra. Se añadió una sonda adicional en la etapa de hibridación: una mezcla 1:4 de sonda de control de hibridación premarcada y no marcada (HYC01). Se usa HYC01 para medir la eficacia de la etapa de hibridación en el portaobjetos y para normalizar las señales de sonda entre portaobjetos. Las micromatrices usadas en el ensayo de lactantes de GA-map se produjeron por Arraylt (Arraylt, Sunnyvale, Estados Unidos). Un portaobjetos de vidrio contiene 24 micromatrices idénticas separadas, y las sondas (complementarias de las sondas enumeradas en la Tabla 5) se aplicaron puntualmente por triplicado en cada matriz. Además, las matrices también incluían dos sondas de control sin unión (NBC01, NBC02) (Sanguin *et al.*, 2006, Environmental Microbiology 8, 289-307).

Tabla 5. Sondas incluidas en el conjunto de sondas 3

ID de sonda	Grupos taxonómicos detectados	Secuencia de sonda	Falso positivo/ falso negativo	SEQ ID NO	Señal correcta media	Señal correcta convencional
1_1	<i>Bacteroides</i>	TTGGGGCTCAACCGTAAATTTG	0 % / 0 %	5	1723,54	245,51
1_1_1_3	<i>Parabacteroides</i>	CGCCTGCCTCAAACATA	0 % / 0 %	19	733,62	N/D
1_2_2	<i>Bacteroides (dorei, fragilis, thetaiotaomicron, vulgatus)</i>	GCACTCAAGACATCCAGTATCAACTG	0 % / 0 %	6	1261,71	435,04
1_3_3	<i>Bacteroides (dorei, fragilis, thetaiotaomicron, vulgatus)</i>	AGGGCAGTCATCCTTCACG	0 % / 0 %	7	1157,96	391,09
2_1_min1b	Gamma-proteobacteria	CAGGTGTAGCGGTGAAATGCGTAGAGAT	14 % / 0 %	20	270,16	N/D
2_1_1_1	<i>Haemophilus</i>	ACGCTCGCAC	0 % / 0 %	21	1711,24	201,24
2_3_2	Subgrupo de Gamma-proteobacteria	CGGGGATTCACATCTGA	8 % / 0 %	22	141,42	N/D
2_4_1	Subgrupo de Gamma-proteobacteria	TGCCAGTTTCGAATGCAGTT	4 % / 0 %	23	1677,81	251,28
2_5_1	Subgrupo de Gamma-proteobacteria	GTGCTTCTCTGCGGGTAA	0 % / 0 %	24	611,51	155,12
2_7_1	<i>Salmonella, Enterobacter, Citrobacter, Cronobacter</i>	TGTTGTGGTTAATAACCGCAGCAATTGA	4 % / 0 %	9	1527,71	N/D
3_2	<i>Proteobacteria</i>	ACGCTTGCACCCCT	5 % / 0 %	1	809,64	278,90
4_1	<i>Firmicutes (Lactobacillales, Clostridium perf, Staphylococcus)</i>	CGATCCGAAAACCTTCTTCACT	6 % / 0 %	2	1799,51	538,14
4_2_3	Subgrupo de <i>Lactobacillus</i>	GCTACACATGGAGTTCCA	29 % / 0 %	25	278,64	14,67
4_3_1	<i>Clostridium ramosum</i>	CCGTCACCTCGGCTACCATTTT	0 % / 0 %	15	2429,10	N/D
4_4_2	<i>Enterococcus, Listeria</i>	TCCAATGACCCCTCCC	0 % / 0 %	10	640,06	125,05
4_5_2	<i>Streptococcus pyogenes</i>	GATTTTCCACTCCACCAT	0 % / 0 %	16	1556,65	N/D
4_6_1	<i>Streptococcus sanguinis</i>	CACTCTCACACCCGTT	0 % / 0 %	11	978,28	N/D
4_7_2	<i>Listeria</i>	CCGTCAAGGGACAAG	0 % / 0 %	17	678,60	N/D
4_8_1	<i>Streptococcus Enterococcus</i>	GTTGCTCGGTCAGACTT	12 % / 0 %	12	1593,28	N/D

ID de sonda	Grupos taxonómicos detectados	Secuencia de sonda	Falso positivo/ falso negativo	SEQ ID NO	Señal correcta media	Señal correcta convencional
5_1	<i>Firmicutes (Clostridia, Bacillales, Enterococcus, Lactobacillus)</i>	GGACAACGCTTGCCAC	6 % / 0 %	3	1315,09	417,36
5_1_2	<i>Staphylococcus</i>	CGTGGCTTTCTGATTAGGTA	0 % / 0 %	13	654,06	N/D
5_2_1	<i>Clostridium neonatale</i>	CGTAGTTAGCCGTGG	0 % / 0 %	26	0,00	0,00
6_2	<i>Bifidobacterium longum</i>	TGCTTATTCAACGGGTAACT	0 % / 0 %	14	2071,50	492,05
6_2	<i>Actinobacteria</i>	CGTAGGCGGTTCGTCGCGT	0 % / 0 %	4	1417,55	243,38
6_2_2	<i>Bifidobacterium, breve</i>	CGGTGCTTATTTCGAAAGGTACACT	0 % / 0 %	18	1928,16	N/D
UNI01	16S Universal	CGTATTACCGCGGCTGCTGGCA	N/D	55	N/D	N/D
HYC01	Control de hibridación	GTAGCATTTCGATTTCGGGCAA	N/D	56	N/D	N/D

Extensión de cebadores e hibridación con la matriz

Antes de la reacción de marcaje los productos de PCR de 16S (amplificados como se ha descrito anteriormente) se trataron con 3 U de exonucleasa I (New England Biolabs, Ipswich, Estados Unidos) y 8 U de fosfatasa alcalina de camarón (USB, Cleveland, Estados Unidos) a 37 °C durante 2 horas y se inactivaron a 80 °C durante 15 min. Los productos de PCR tratados con ExoSAP se cuantificaron después usando software de captura de imágenes moleculares Kodak (Versión 4.0) basándose en imágenes de electroforesis en gel. Se incluyó en todos los geles una escalera de ADN de 1 kB (N3232, New England Biolabs) con concentraciones especificadas. Basándose en la cuantificación de las imágenes de gel los productos de PCR se diluyeron hasta una concentración igual de 50 ng/μl por muestra y se usaron aproximadamente 100 ng de molde en la siguiente reacción de marcaje: en un volumen de reacción total de 10 μl de 2,5 U de HOT TERMIPOl (Solis Biodyne), tampón C 1x (Solis Biodyne), MgCl₂ 4 mM (Solis Biodyne), ddCTP-tamra 0,4 μM (Jena Bioscience, Jena, Alemania) y conjunto de sondas 3 2,9 μM (Tabla 5). El protocolo de marcaje incluyó un estadio de activación de 12 min a 95 °C, seguido de 10 ciclos con 20 s de desnaturalización a 96 °C y 35 s de hibridación a 60 °C. Se seleccionaron aleatoriamente cuarenta y cuatro muestras para examinar la reproducibilidad. Estas 44 muestras se procesaron dos veces comenzando a partir de la reacción de marcaje. Además, como un ensayo del intervalo cuantitativo de los productos de PCR de ensayo a partir de cultivos puros de 5 especies diferentes (enumerado en la Tabla 4) se diluyeron de 10⁰ a 10⁻⁴ y se incluyeron en la reacción de marcaje y análisis de matriz corriente abajo.

Las matrices se prehibridaron para evitar la señal de fondo empapando los portaobjetos de vidrio en BlockIt (ArrayIt) a temperatura ambiente. Después de dos horas los portaobjetos se lavaron durante 2 min en un tampón de lavado que contenía SSC 2x (Sigma-Aldrich, St. Louis, Estados Unidos) + sarcosilo 0,1 % (TA) (VWR, International Ltd., Poole, Reino Unido) y después durante 2 min en SSC 2x (Sigma-Aldrich). Los portaobjetos se colocaron después en un vaso de precipitados con H₂O ultra pura (100 °C) durante 2 min y se transfirieron inmediatamente a un vaso de precipitados que contenía etanol 100 % (-20 °C) durante 20 s, antes de secarse por centrifugación a 91 G en una centrifuga Multifuge 3 S-R (Heraeus, Buckinghamshire, Reino Unido) durante 12 min y se usaron antes de una hora.

Inmediatamente antes de la propia hibridación de matriz se añadieron a las muestras 60 μl de tampón de hibridación que contenía polietilenglicol 800 7,2 % (Sigma-Aldrich), SSC 1,2x (Sigma-Aldrich) y 0,17 μM de la mezcla de sonda de control de hibridación HYC01 (mezcla 1:4 de HYC01 marcado con tamra y HYC01 no marcado). Las muestras se desnaturalizaron a 95 °C durante 2 min y después se dejaron a 45 °C durante 2 min. Los portaobjetos de vidrio se colocaron en una cámara de hibridación de 96 pocillos (ArrayIt) antes de cargarse las muestras en las matrices. Se usaron dos matrices por portaobjeto para las muestras de control positivas y negativas. La cámara de hibridación se colocó en una cámara húmeda y se hibridó durante 16 horas en un agitador incubador Innova 4000 (New Brunswick Scientific, Champaign, Estados Unidos) a 45 °C y 60 rpm.

Después de la hibridación las matrices se lavaron durante 5 minutos en el tampón de lavado que contenía SSC 2x (Sigma-Aldrich) y sarcosilo 0,1 % (VWR, International Ltd.), después durante 5 min en SSC 2x (Sigma-Aldrich) y finalmente durante 10 s en SSC 0,2x (Sigma-Aldrich), antes de secarse por centrifugación a 91 G durante 12 min en una centrifuga Multifuge 3 S-R (Heraeus). Se exploraron matrices hibridadas a una longitud de onda de 532 nm con un explorador recargado Tecan LS (Tecan, Männedorf, Austria). Se analizaron las intensidades fluorescentes y morfologías puntuales usando Axon GenePix Pro 6.0.

Electroforesis capilar

Para ensayar la especificidad de sonda, se ensayaron sondas individuales contra sus bacterias diana (ADN de cultivos puros) realizando amplificación por PCR 16S y reacciones de marcaje como se ha descrito anteriormente (con sondas individuales 1 μM en lugar del conjunto de sondas 3) y el rendimiento de las sondas se evaluó usando electroforesis capilar. Se examinó la reproducibilidad de la PRC del gen de ARNr 16S en una de las muestras (amplificada en tres reacciones de PCR separadas) usando electroforesis capilar. Se eligieron dos sondas (6_1_4 y 5_1_2) para examinar la señal para cada uno de los tres productos de PCR, y también se examinó un procesamiento por triplicado en un grupo de los tres productos de PCR usando las mismas sondas. Después del marcaje, las muestras se trataron con 8 U de SAP (USB) y se incubaron a 37 °C durante 1 hora y se inactivaron a 80 °C durante 15 min. Después se mezclaron 1 μl de las sondas tratadas con SAP y marcadas con 9 μl de formamida Hi-Di (Applied Biosystems, Warrington, Reino Unido) y 0,5 μl de patrón de tamaño GeneScan 120 Liz (Applied Biosystems), y las muestras se incubaron a 95 °C durante 5 min, y se pusieron en hielo inmediatamente. Las muestras se cargaron después en una matriz capilar 3130xl de 50 cm (Applied Biosystems) en el secuenciador ABI Genetic Analyzer 3130xl (Applied Biosystems), que contenía el polímero de rendimiento optimizado 7 (POP-7, Applied Biosystems). El tiempo de inyección fue de 16-22 s y las condiciones electroforéticas fueron: tiempo de procesamiento 1.500 s a 15.000 V, corriente de procesamiento 100 μA y temperatura de procesamiento 60 °C. Se usó el software GeneMapper 4.0 para analizar los resultados.

Los productos de PCR de gen de ARN 16S de las 26 cepas bacterianas usados para evaluar las sondas se secuenciaron para confirmar su identidad y examinar si hubo cualquier mutación en sus secuencias génicas en comparación con las secuencias usadas para diseñar las sondas. Los productos de PCR tratados con exoSAP se diluyeron 10 veces y se usó 1 μl en la reacción de secuenciación usando el kit de secuenciación en ciclo BigDye®

Terminator v1.1 (Applied Biosystems). Se usaron 0,32 μ M de los mismos cebadores directos e inversos usados para la PCR de ARNr 16S descrita anteriormente en dos reacciones de secuenciación separadas. Se usó el kit de purificación BigDye XTerminator® (Applied Biosystems, Warrington, Reino Unido) de acuerdo con las recomendaciones del fabricante para limpiar las reacciones de secuenciación. Las muestras se analizaron en una matriz capilar 3130xl de 36 cm (Applied Biosystems) en el secuenciador ABI Genetic Analyzer 3130xl (Applied Biosystems), que contenía el polímero de rendimiento optimizado 7 (POP-7, Applied Biosystems). El tiempo de inyección fue de 3 s y las condiciones fueron: tiempo de procesamiento 2780 s a 8.500 V, corriente de procesamiento 5,0 μ A y temperatura de procesamiento de 60 °C. Se determinaron las bases de las secuencias por el software explorador de secuencias v1.0 (Applied Biosystems). Las secuencias se han depositado en GenBank y los números de referencia respectivos de las cepas se enumeran en la Tabla 4.

Preprocesamiento y análisis de datos

Las señales de sonda se corrigieron con respecto a variaciones de hibridación no deseadas que se observan entre portaobjetos. En cada experimento, se añade una sonda que ya está marcada (HYC01) a la mezcla de sondas para evaluar la etapa de hibridación. Para corregir con respecto a hibridación variante entre portaobjetos, se dividieron todas las señales de muestra en la señal promedio de todas las réplicas de esta sonda. Además, se retiró la señal de fondo de cada sonda individual restando las señales promedio de una muestra de control negativo incluida en todos los portaobjetos usados en este experimento.

Se construyó un árbol de unión de vecinos de las 26 cepas bacterianas para evaluar las sondas usando el programa Mega 4 y su fiabilidad se infirió usando bootstrapping.

Análisis estadísticos

La especificidad de sonda se evaluó comparando los valores diana / no diana teóricos con los resultados experimentales en cepas individuales, usando un valor de umbral de señal de fondo determinado de forma empírica de 50.

Los datos de micromatrices habitualmente contienen valores tanto de umbral como de saturación y por lo tanto se distribuyen de forma normal muy pocas veces. Por lo tanto, para ensayar la significación de datos de micromatrices, es habitual usar enfoques basados en permutación en lugar de ensayos estadísticos convencionales tales como ANOVA y ensayos de t (que requieren distribución normal). El ensayo de permutación es un ensayo estadístico exacto, incluso para datos con una estructura de distribución compleja. Por lo tanto, los p valores para diferencias de grupo dentro de cada categoría de edad se calcularon por ensayo de permutación (Langsrud, 2002, Journal Of The Royal Statistical Society Series D 51,305-317) usando 50 como el valor umbral de fondo.

Resultados

Construcción y evaluación de sondas

Se construyó un conjunto de 88 sondas basándose en los criterios descritos en Materiales y Métodos. Seis sondas para los filos principales abarcaron 88 % de los clones en el conjunto de datos de los inventores. Las evaluaciones de sondas individuales usando electroforesis de gel capilar y las cepas en la Tabla 4 como moldes mostraron que el 76 % de las sondas satisfacían el criterio de detección de diana, lo que indica una tasa de éxito relativamente alta. Basándose en estos resultados y un conjunto de criterios bioinformáticos se identificaron 10 conjuntos de sondas. Cada conjunto de sondas consistía en 25 sondas que se seleccionaron basándose en su compatibilidad informática entre sí. Las estimaciones de compatibilidad se basaron en cálculos de la temperatura de fusión y termodinámica de la sonda, autohibridación, hibridación con otras sondas en el conjunto de sondas o sus bacterias diana. La validación experimental por electroforesis de gel capilar mostró que el conjunto de sondas 3 proporcionó la reacción cruzada más baja, como se determina marcando sin molde (resultados no mostrados). Este conjunto de sondas se seleccionó por lo tanto para construcción de matrices (Tabla 5).

Especificidad, reproducibilidad e intervalo cuantitativo de la matriz de lactantes de GA-map

La primera evaluación de la matriz fue en cepas puras. La evaluación se basó en la comparación por ordenador de determinadas dianas / no dianas con la de señales experimentales. Este análisis mostró buena concordancia entre las especificidades de sonda teóricas y experimentales. Usando un valor de punto de corte de señal de 50 ufr se descubrió que no hubo falsos negativos, mientras que el número de falsos positivos era más bien variable (Tabla 5).

La siguiente etapa en la evaluación fue determinar la precisión de clasificación de muestras mixtas. Esto se realizó analizando un conjunto de mezclas definidas. La evolución de estos datos mostró que la mayoría de las sondas identificaban con precisión sus bacterias diana. Aquí, se usó una señal de umbral de 50. En total, hubo 96,5 % de correctos con 9,0 % de falsos positivos y 1,6 % de falsos negativos. El intervalo cuantitativo de las sondas seleccionadas se evaluó posteriormente por diluciones de molde. En general, estos análisis mostraron que hubo una saturación de la señal de sonda cuando la concentración diana fue > 10 % en relación con el producto de PCR no

diluido. Todas las sondas evaluadas también proporcionaron el mismo límite de detección aproximado entre 0,1 y 0,01 %. La precisión cuantitativa también fue muy alta con un $R^2 > 0,9$ para todas las sondas ensayadas. La reproducibilidad del ensayo se evaluó por análisis duplicado de 43 muestras. Se determinó la variación de porcentaje media y R^2 para cada sonda individualmente y se confirmó la reproducibilidad del ensayo.

Desarrollo en el nivel de filo de la microbiota del intestino

Se descubrió que las actinobacterias (sonda 6_2) y firmicutes (sonda 5_1) estaban significativamente sobrerrepresentadas a los 4 meses y un año, respectivamente, en los niños sensibilizados con IgE (Tabla 6 y Figura 1). También hubo un patrón de colonización específico de la edad uniforme general en el nivel de filo, independientemente del estado de sensibilización. El patrón general fue una dominancia inicial de firmicutes y proteobacterias a los diez días. A los cuatro meses la dominancia de proteobacterias / firmicutes se reemplazó con bacteroides / actinobacterias, mientras que después de uno y dos años aparentemente se estaba reduciendo la abundancia de filos colonizantes iniciales.

Tabla 6. Diferencias en el nivel de filo entre niños sensibilizados y no sensibilizados a diferentes edades¹

Sonda	10 días	120 días	360 días	720 días
1_1	0,640	0,868	1,00	0,903
2_1_min1b	0,760	0,220	0,801	0,542
3_2	0,922	0,3126	0,126	0,465
4_1	0,164	0,190	0,360	0,599
5_1	0,486	0,127	0,049	0,556
6_2	0,152	0,042	0,196	0,989
UNI01	0,450	0,867	0,917	0,216

¹ La significación de las diferencias se determinó por ensayo de permutación. Las diferencias significativas ($p < 0,05$) están en negrita.

Desarrollo de la microbiota del intestino en el nivel de género y especie

La diferencia principal entre el grupo sensibilizado y no sensibilizado fue que *B. longum* (sonda 6_1_4) estaba significativamente sobrerrepresentado en el grupo sensibilizado, en comparación con el grupo no sensibilizado en un año. También se descubrió que *Enterococcus* (sonda 4_4_2) estaba significativamente sobrerrepresentado a los cuatro meses. También parece que los estreptococos se asocian con la sensibilización, estando el *Streptococcus sanguinis* (sonda 4_6_1) significativamente sobrerrepresentado en un año, y *S. pneumonia* (sonda 4_8_1) en límite de la significación a los 10 días (Tabla 7 y Figura 2).

Los grupos bacterianos con los patrones de colonización más uniformes que se correlacionaban con la edad eran *Staphylococcus* (sonda 5_1_2) y *Bifidobacterium breve* (sonda 6_2_2). El *Staphylococcus* dominó inicialmente, mientras que *B. breve* tuvo un pico de dominancia a los 4 meses.

Tabla 7. Diferencias de género/ especie entre niños sensibilizados y no sensibilizados a diferentes edades

Sonda	10 días	120 días	360 días	720 días
1_1_3	1	0,866	1,000	1,000
1_2_2	1	0,884	1,000	1,000
1_3_3	0,756	0,488	0,206	0,741
2_1_1	0,783	1,000	1,000	1,000
2_3_2	0,668	0,347	1,000	0,494
2_4_1	0,182	0,622	1,000	1,000
2_5_1	0,695	0,913	0,870	0,949
2_7_1	0,754	1,000	1,000	1,000
4_2_3	0,938	0,909	1,000	0,405
4_3_1	0,786	0,765	0,828	0,537

Sonda	10 días	120 días	360 días	720 días
4_4_2	0,9736	0,020	1,000	1,000
4_6_1	1,000	1,000	0,038	0,689
4_8_1	<i>0,084</i>	0,169	1,000	0,935
5_1_2	0,847	1,000	1,000	0,399
6_1_4	<i>0,097</i>	<i>0,066</i>	0,016	0,837
6_2_1	0,933	0,741	0,863	0,857
6_2_2	0,711	0,679	0,844	0,784

[†] La significación de las diferencias se determinó por ensayo de permutación. Las diferencias significativas ($p < 0.05$) están en negrita, mientras que las diferencias en el intervalo de $0,05 < p < 0,1$ están en cursiva.

Análisis

5 Con el ensayo de GA-map basado en SNUPE se ha obtenido una micromatriz de genes de ARNr 16S de alta especificidad y sensibilidad con solamente algunas sondas. El beneficio evidente de esto es que el ensayo permite aplicaciones de alto rendimiento.

10 El hallazgo biológico más sorprendente en los datos de los inventores fue que *B. longum* estaba significativamente sobrerrepresentado en el grupo sensibilizado con IgE a los 360 días, además de p valores bajos para 10 días y 120 días. Este hallazgo también se ha confirmado independientemente por q-PCR para los datos de IM-PACT (resultados no publicados). Tomadas juntas, las correlaciones independientes múltiples apoyan la validez de las observaciones. La sorpresa fue porque la mayoría del trabajo previo ha sugerido de hecho que *B. longum* es protector con respecto a sensibilización. Los experimentos con modelos de ratón, sin embargo, han mostrado que el momento y orden de colonización por bifidobacterias son importantes para los efectos inmunomoduladores. Esto puede explicar las diferencias en los efectos entre diferentes estudios.

20 También se descubrió que el subgrupo de firmicutes que contenía estreptococos y enterococos estaba significativamente sobrerrepresentado para el grupo sensibilizado con IgE. Se ha descrito relativamente poco acerca de estos grupos bacterianos con respecto a la sensibilización. Se ha sugerido, sin embargo, que las infecciones por *S. pneumonia* pueden correlacionarse con niveles de IgE aumentados en bronquitis crónica. Por lo tanto, podría haber mecanismos subyacentes comunes para la sensibilización del lactante y bronquitis.

25 En resumen, este estudio demuestra la utilidad del ensayo GA-map del lactante en la determinación de variaciones en la composición de la microbiota del intestino. Dicha información podría conducir a diagnóstico temprano de la enfermedad y mejores tratamientos profilácticos o terapéuticos de diversas enfermedades relacionadas con el intestino.

LISTADO DE SECUENCIAS

30 <110> Genetic Analysis

<120> Conjunto de sondas de oligonucleótidos y métodos para realizar el perfil de microbiota

35 <130> 463.10776201

<140> Aún no asignado

<141> 16-12-2011

40 <150> GB 1021399.9

<151> 16-12-2010

<160> 56

45 <170> PatentIn versión 3.5

<210> 1

<211> 13

<212> ADN

50 <213> Secuencia Artificial

	<220> <223> Sonda	
5	<400> 1 acgcttgac cct	13
10	<210> 2 <211> 22 <212> ADN <213> Secuencia Artificial	
15	<220> <223> Sonda	
15	<400> 2 cgatccgaaa accttctca ct	22
20	<210> 3 <211> 16 <212> ADN <213> Secuencia Artificial	
25	<220> <223> Sonda	
25	<400> 3 ggacaacgct tgccac	16
30	<210> 4 <211> 19 <212> ADN <213> Secuencia Artificial	
35	<220> <223> Sonda	
40	<400> 4 cgtaggcggt tcgtcggt	19
40	<210> 5 <211> 22 <212> ADN <213> Secuencia Artificial	
45	<220> <223> Sonda	
50	<400> 5 ttgcggctca accgtaaaat tg	22
55	<210> 6 <211> 26 <212> ADN <213> Secuencia Artificial	
55	<220> <223> Sonda	
60	<400> 6 gcactcaaga catccagtat caactg	26
65	<210> 7 <211> 19 <212> ADN <213> Secuencia Artificial	

	<220> <223> Sonda	
5	<400> 7 agggcagtc tcttcacg	19
10	<210> 8 <211> 27 <212> ADN <213> Secuencia Artificial	
15	<220> <223> Sonda	
15	<400> 8 caatcggagt tcttcgtgat atctaag	27
20	<210> 9 <211> 28 <212> ADN <213> Secuencia Artificial	
25	<220> <223> Sonda	
25	<400> 9 tgtgtggtt aataaccgca gcaattga	28
30	<210> 10 <211> 15 <212> ADN <213> Secuencia Artificial	
35	<220> <223> Sonda	
40	<400> 10 tccaatgacc ctccc	15
40	<210> 11 <211> 16 <212> ADN <213> Secuencia Artificial	
45	<220> <223> Sonda	
50	<400> 11 cactctcaca cccggt	16
50	<210> 12	
55	<211> 17 <212> ADN <213> Secuencia Artificial	
55	<220> <223> Sonda	
60	<400> 12 gttgctcgtt cagactt	17
65	<210> 13 <211> 20 <212> ADN <213> Secuencia Artificial	

	<220> <223> Sonda	
5	<400> 13 cgtaggtttc tgattagga	20
10	<210> 14 <211>21 <212> ADN <213> Secuencia Artificial	
15	<220> <223> Sonda	
20	<400> 14 tgcttattca acgggtaaac t	21
25	<210> 15 <211>21 <212> ADN <213> Secuencia Artificial	
30	<220> <223> Sonda	
35	<400> 15 ccgtcactcg gctaccattt c	21
40	<210> 16 <211> 19 <212> ADN <213> Secuencia Artificial	
45	<220> <223> Sonda	
50	<400> 16 gattttccac tcccaccat	19
55	<210> 17 <211> 15 <212> ADN <213> Secuencia artificial	
60	<220> <223> Sonda	
65	<400> 17 ccgtcaaggg acaag	15
	<210> 18 <211> 24 <212> ADN <213> Secuencia Artificial	
	<220> <223> Sonda	
	<400> 18 cggtagcttat tcgaaaggta cact	24
	<210> 19 <211> 17 <212> ADN <213> Secuencia Artificial	

	<220> <223> Sonda	
5	<400> 19 cgctgcctc aaacata	17
10	<210> 20 <211> 28 <212> ADN <213> Secuencia Artificial	
15	<220> <223> Sonda	
15	<400> 20 caggtgtagc ggtgaaatgc gtagagat	28
20	<210> 21 <211> 11 <212> ADN <213> Secuencia Artificial	
25	<220> <223> Sonda	
25	<400> 21 acgctgcac c	11
30	<210> 22 <211> 18 <212> ADN <213> Secuencia Artificial	
35	<220> <223> Sonda	
40	<400> 22 cggggatttc acatctga	18
45	<210> 23 <211> 20 <212> ADN <213> Secuencia Artificial	
50	<220> <223> Sonda	
50	<400> 23 tgccagtttc gaatgcagtt	20
55	<210> 24 <211> 19 <212> ADN <213> Secuencia Artificial	
60	<220> <223> Sonda	
60	<400> 24 gtgcttcttc tgcgggtaa	19
65	<210> 25 <211> 18 <212> ADN <213> Secuencia Artificial	

	<220> <223> Sonda	
5	<400> 25 gctacacatg gagttcca	18
10	<210> 26 <211> 15 <212> ADN <213> Secuencia Artificial	
15	<220> <223> Sonda	
15	<400> 26 cgtagttagc cgtgg	15
20	<210> 27 <211> 13 <212> ADN <213> Secuencia Artificial	
25	<220> <223> Sonda	
25	<400> 27 agggtgcaag cgt	13
30	<210> 28 <211> 22 <212> ADN <213> Secuencia Artificial	
35	<220> <223> Sonda	
40	<400> 28 agtgaagaag gtttcggat cg	22
40	<210> 29 <211> 16 <212> ADN <213> Secuencia Artificial	
45	<220> <223> Sonda	
50	<400> 29 gtggcaagcg ttgtcc	16
55	<210> 30 <211> 19 <212> ADN <213> Secuencia Artificial	
55	<220> <223> Sonda	
60	<400> 30 acgcgacgaa ccgcctacg	19
65	<210> 31 <211> 22 <212> ADN <213> Secuencia Artificial	

	<220>	
	<223> Sonda	
5	<400> 31 caattttacg gttgagccgc aa	22
	<210> 32	
	<211> 26	
10	<212> ADN	
	<213> Secuencia Artificial	
	<220>	
	<223> Sonda	
15	<400> 32 cagttgatac tggatgtctt gagtgc	26
	<210> 33	
	<211> 19	
20	<212> ADN	
	<213> Secuencia Artificial	
	<220>	
	<223> Sonda	
25	<400> 33 cgtgaaggat gactgccct	19
	<210> 34	
30	<211> 27	
	<212> ADN	
	<213> Secuencia Artificial	
	<220>	
35	<223> Sonda	
	<400> 34 cttagatatac acgaagaact ccgattg	27
40	<210> 35	
	<211> 28	
	<212> ADN	
	<213> Secuencia Artificial	
45	<220>	
	<223> Sonda	
	<400> 35 tcaattgctg cgggtattaa ccacaaca	28
50	<210> 36	
	<211> 15	
	<212> ADN	
	<213> Secuencia Artificial	
55	<220>	
	<223> Sonda	
	<400> 36 gggaggtca ttgga	15
	<210> 37	
	<211> 16	
	<212> ADN	
65	<213> Secuencia Artificial	

	<220> <223> Sonda	
5	<400> 37 aacgggtgtg agagtg	16
10	<210> 38 <211> 17 <212> ADN <213> Secuencia Artificial	
15	<220> <223> Sonda	
20	<400> 38 aagtctgacc gagcaac	17
25	<210> 39 <211> 20 <212> ADN <213> Secuencia Artificial	
30	<220> <223> Sonda	
35	<400> 39 tacctaataca gaaagccacg	20
40	<210> 40 <211> 21 <212> ADN <213> Secuencia Artificial	
45	<220> <223> Sonda	
50	<400> 40 agtttaccg ttgaataagc a	21
55	<210> 41 <211> 21 <212> ADN <213> Secuencia Artificial	
60	<220> <223> Sonda	
65	<400> 41 gaaatggtag ccgagtgacg g	21
	<210> 42 <211> 19 <212> ADN <213> Secuencia Artificial	
	<220> <223> Sonda	
	<400> 42 atggtgggag tggaaaatc	19
	<210> 43 <211> 15 <212> ADN <213> Secuencia Artificial	

	<220> <223> Sonda	
5	<400> 43 ctgtccctt gacgg	15
10	<210> 44 <211> 24 <212> ADN <213> Secuencia Artificial	
15	<220> <223> Sonda	
20	<400> 44 agtgtacctt tcgaataagc accg	24
25	<210> 45 <211> 17 <212> ADN <213> Secuencia Artificial	
30	<220> <223> Sonda	
35	<400> 45 tatgtttgag gcaggcg	17
40	<210> 46 <211> 28 <212> ADN <213> Secuencia Artificial	
45	<220> <223> Sonda	
50	<400> 46 atctctacgc atttcaccgc tacacctg	28
55	<210> 47 <211> 11 <212> ADN <213> Secuencia Artificial	
60	<220> <223> Sonda	
65	<400> 47 ggtgcgagcg t	11
	<210> 48 <211> 18 <212> ADN <213> Secuencia Artificial	
	<220> <223> Sonda	
	<400> 48 tcagatgtga aatccccg	18
	<210> 49	
	<211> 20 <212> ADN <213> Secuencia Artificial	

	<220> <223> Sonda	
5	<400> 49 aactgcattc gaaactggca	20
10	<210> 50 <211> 19 <212> ADN <213> Secuencia Artificial	
15	<220> <223> Sonda	
20	<400> 50 ttacccgcag aagaagcac	19
25	<210> 51 <211> 18 <212> ADN <213> Secuencia Artificial	
30	<220> <223> Sonda	
35	<400> 51 tggaactcca tgtgtagc	18
40	<210> 52 <211> 15 <212> ADN <213> Secuencia Artificial	
45	<220> <223> Sonda	
50	<400> 52 ccacggctaa ctacg	15
55	<210> 53 <211> 18 <212> ADN <213> Secuencia Artificial	
60	<220> <223> Sonda	
65	<400> 53 tcctacggga ggcagcag	18
	<210> 54 <211> 20 <212> ADN <213> Secuencia artificial	
	<220> <223> Sonda	
	<400> 54 cggttacctt gttacgactt	20
	<210> 55 <211> 22 <212> ADN <213> Secuencia Artificial	

ES 2 573 133 T3

	<220>	
	<223> Sonda	
5	<400> 55 cgtattaccg cggctgctgg ca	22
	<210> 56	
	<211> 20	
10	<212> ADN	
	<213> Secuencia Artificial	
	<220>	
	<223> Sonda	
15	<400> 56 gtagcattcg attcgggcaa	20

REIVINDICACIONES

1. Un conjunto de sondas oligonucleotídicas, comprendiendo dicho conjunto:

- 5 (a) un oligonucleótido que consiste en (ai) una secuencia de nucleótidos seleccionada de ACGCTTGACCCCT (SEQ ID NO 1), opcionalmente con hasta tres bases sustituidas, y AGGGTGCAAGCGT (SEQ ID NO 27), opcionalmente con hasta tres bases sustituidas, y (aii) 0-5 nucleótidos además de (ai), en el que el oligonucleótido (a) es capaz de hibridar con una SEQ ID NO 27 o SEQ ID NO 1, respectivamente, en condiciones de alta rigurosidad;
- 10 (b) un oligonucleótido que consiste en (bi) una secuencia de nucleótidos seleccionada de CGATCCGAAAACCTTCTTCACT (SEQ ID NO 2), opcionalmente con hasta tres bases sustituidas, y AGTGAAGAAGGTTTTCGGATCG (SEQ ID NO 28), opcionalmente con hasta tres bases sustituidas, y (bii) 0-5 nucleótidos además de (bi), en el que el oligonucleótido (b) es capaz de hibridar con una SEQ ID NO 28 o SEQ ID NO 2, respectivamente, en condiciones de alta rigurosidad;
- 15 (c) un oligonucleótido que consiste en (ci) una secuencia de nucleótidos seleccionada de GGACAACGCTTGCCAC (SEQ ID NO 3), opcionalmente con hasta tres bases sustituidas, y GTGGCAAGCGTTGTCC (SEQ ID NO 29), opcionalmente con hasta tres bases sustituidas, y (cii) 0-5 nucleótidos además de (ci), en el que el oligonucleótido (c) es capaz de hibridar con una SEQ ID NO 29 o SEQ ID NO 3, respectivamente, en condiciones de alta rigurosidad;
- 20 (d) un oligonucleótido que consiste en (di) una secuencia de nucleótidos seleccionada de CGTAGGCGGTTCGTCGCGT (SEQ ID NO 4), opcionalmente con hasta tres bases sustituidas, y ACGCGACGAACCGCCTACG (SEQ ID NO 30), opcionalmente con hasta tres bases sustituidas, y (dii) 0-5 nucleótidos además de (di), en el que el oligonucleótido (d) es capaz de hibridar con SEQ ID NO 30 o SEQ ID NO 4, respectivamente, en condiciones de alta rigurosidad; y
- 25 (e) uno o más oligonucleótidos que consisten en (ei) una secuencia de nucleótidos seleccionada de las enumeradas como SEQ ID NO 5-8 y 31-34 en la Tabla 1, opcionalmente con hasta tres bases sustituidas, y (eii) 0-5 nucleótidos además de (ei), en el que el oligonucleótido (e) es capaz de hibridar con la secuencia de nucleótidos complementaria a SEQ ID NO 5-8 y 31-34, respectivamente, en condiciones de alta rigurosidad; y opcionalmente
- 30 (f) uno o más oligonucleótidos que consisten en (fi) una secuencia de nucleótidos seleccionada de las enumeradas como SEQ ID NO 9-26 y 35-52 en la Tabla 2 y Tabla 3, opcionalmente con hasta tres bases sustituidas y (fii) 0-5 nucleótidos además de (fi), en el que el oligonucleótido (f) es capaz de hibridar con la secuencia de nucleótidos complementaria a SEQ ID NO 9-26 y 35-52, respectivamente, en condiciones de alta rigurosidad.

2. El conjunto de sondas oligonucleotídicas de la reivindicación 1, comprendiendo dicho conjunto:

- 35 (a) un oligonucleótido que consiste en (ai) la secuencia de nucleótidos de SEQ ID NO 1 o SEQ ID NO 27, y (aii) 0-5 nucleótidos además de (ai), en el que el oligonucleótido (a) es capaz de hibridar con SEQ ID NO 27 o SEQ ID NO 1, respectivamente, en condiciones de alta rigurosidad;
- (b) un oligonucleótido que consiste en (bi) la secuencia de nucleótidos de SEQ ID NO 2 o SEQ ID NO 28, y (bii) 0-5 nucleótidos además de (bi), en el que el oligonucleótido (b) es capaz de hibridar con SEQ ID NO 28 o SEQ ID NO 2, respectivamente, en condiciones de alta rigurosidad;
- 40 (c) un oligonucleótido que consiste en (ci) la secuencia de nucleótidos de SEQ ID NO 3 o SEQ ID NO 29, y (cii) 0-5 nucleótidos además de (ci), en el que el oligonucleótido (c) es capaz de hibridar con SEQ ID NO 29 o SEQ ID NO 3, respectivamente, en condiciones de alta rigurosidad;
- (d) un oligonucleótido que consiste en (di) la secuencia de nucleótidos de SEQ ID NO 4 o SEQ ID NO 30, y (dii) 0-5 nucleótidos además de (di), en el que el oligonucleótido (d) es capaz de hibridar con SEQ ID NO 30 o SEQ ID NO 4, respectivamente, en condiciones de alta rigurosidad; y
- 45 (e) uno o más oligonucleótidos que consisten en (ei) una secuencia de nucleótidos seleccionada de las enumeradas como SEQ ID NO 5-8 y 31-34 en la Tabla 1 y (eii) 0-5 nucleótidos además de (ei), en el que el oligonucleótido (e) es capaz de hibridar con la secuencia de nucleótidos complementaria a SEQ ID NO 5-8 y 31-34, respectivamente, en condiciones de alta rigurosidad; y opcionalmente
- 50 (f) uno o más oligonucleótidos que consisten en (fi) una secuencia de nucleótidos seleccionada de las enumeradas como SEQ ID NO 9-26 y 35-52 en la Tabla 2 y Tabla 3 y (fii) 0-5 nucleótidos además de (fi), en el que el oligonucleótido (f) es capaz de hibridar con la secuencia de nucleótidos complementaria a SEQ ID NO 9-26 y 35-52, respectivamente, en condiciones de alta rigurosidad.

3. El conjunto de sondas oligonucleotídicas de la reivindicación 1, comprendiendo dicho conjunto:

- 60 (a) un oligonucleótido que consiste en una secuencia de nucleótidos seleccionada de SEQ ID NO 1, opcionalmente con hasta tres bases sustituidas, y SEQ ID NO 27, opcionalmente con hasta tres bases sustituidas, en el que el oligonucleótido (a) es capaz de hibridar con SEQ ID NO 27 o SEQ ID NO 1, respectivamente, en condiciones de alta rigurosidad;
- (b) un oligonucleótido que consiste en una secuencia de nucleótidos seleccionada de SEQ ID NO 2, opcionalmente con hasta tres bases sustituidas, y SEQ ID NO 28, opcionalmente con hasta tres bases sustituidas, en el que el oligonucleótido (b) es capaz de hibridar con SEQ ID NO 28 o SEQ ID NO 2, respectivamente, en condiciones de alta rigurosidad;
- 65

(c) un oligonucleótido que consiste en una secuencia de nucleótidos seleccionada de SEQ ID NO 3, opcionalmente con hasta tres bases sustituidas, y SEQ ID NO 29, opcionalmente con hasta tres bases sustituidas, en el que el oligonucleótido (c) es capaz de hibridar con SEQ ID NO 29 o SEQ ID NO 3, respectivamente, en condiciones de alta rigurosidad;

(d) un oligonucleótido que consiste en una secuencia de nucleótidos seleccionada de SEQ ID NO 4, opcionalmente con hasta tres bases sustituidas, y SEQ ID NO 30, opcionalmente con hasta tres bases sustituidas, en el que el oligonucleótido (d) es capaz de hibridar con SEQ ID NO 30 o SEQ ID NO 4, respectivamente, en condiciones de alta rigurosidad; y

(e) uno o más oligonucleótidos que consisten en una secuencia de nucleótidos seleccionada de las enumeradas como SEQ ID NO 5-8 y 31-34 en la Tabla 1, opcionalmente con hasta tres bases sustituidas, en el que el oligonucleótido (e) es capaz de hibridar con la secuencia de nucleótidos complementaria a SEQ ID NO 5-8 y 31-34, respectivamente, en condiciones de alta rigurosidad; y opcionalmente

(f) uno o más oligonucleótidos que consisten en una secuencia de nucleótidos seleccionada de las enumeradas como SEQ ID NO 9-26 y 35-52 en la Tabla 2 y Tabla 3, opcionalmente con hasta tres bases sustituidas, en el que el oligonucleótido (f) es capaz de hibridar con la secuencia de nucleótidos complementaria a SEQ ID NO 9-26 y 35-52, respectivamente, en condiciones de alta rigurosidad.

4. El conjunto de sondas oligonucleotídicas de la reivindicación 1, comprendiendo dicho conjunto:

(a) un oligonucleótido que consiste en la secuencia de nucleótidos seleccionada de SEQ ID NO 1, SEQ ID NO 1 con un resto C 3' adicional, SEQ ID NO 27 o SEQ ID NO 27 con un resto G 5' adicional;

(b) un oligonucleótido que consiste en la secuencia de nucleótidos de SEQ ID NO 2, SEQ ID NO 2 con un resto C 3' adicional, SEQ ID NO 28 o SEQ ID NO 28 con un resto G 5' adicional;

(c) un oligonucleótido que consiste en la secuencia de nucleótidos de SEQ ID NO 3, SEQ ID NO 3 con un resto C 3' adicional, SEQ ID NO 29 o SEQ ID NO 29 con un resto G 5' adicional;

(d) un oligonucleótido que consiste en la secuencia de nucleótidos de SEQ ID NO 4, SEQ ID NO 4 con un resto C 3' adicional, SEQ ID NO 30 o SEQ ID NO 30 con un resto G 5' adicional; y

(e) uno o más oligonucleótidos que consisten en una secuencia de nucleótidos seleccionada de las enumeradas en la Tabla 1 como SEQ ID NO 5-8 y SEQ ID NO 31-34, o SEQ ID NO 5-8 con un resto C 3' adicional o SEQ ID NO 31-34 con un resto G 5' adicional; y opcionalmente

(f) uno o más oligonucleótidos que consisten en una secuencia de nucleótidos seleccionada de las enumeradas en la Tabla 2 y Tabla 3 como SEQ ID NO 9-26 y SEQ ID NO 35-52, o SEQ ID NO 9-26 con un resto C 3' adicional, o SEQ ID NO 35-52 con un resto G 5' adicional.

5. El conjunto de sondas oligonucleotídicas de la reivindicación 1, comprendiendo dicho conjunto:

(a) un oligonucleótido que consiste en la secuencia de nucleótidos de SEQ ID NO 1 o SEQ ID NO 27;

(b) un oligonucleótido que consiste en la secuencia de nucleótidos de SEQ ID NO 2 o SEQ ID NO 28;

(c) un oligonucleótido que consiste en la secuencia de nucleótidos de SEQ ID NO 3 o SEQ ID NO 29;

(d) un oligonucleótido que consiste en la secuencia de nucleótidos de SEQ ID NO 4 o SEQ ID NO 30; y

(e) uno o más oligonucleótidos que consisten en una secuencia de nucleótidos seleccionada de las enumeradas en la Tabla 1; y opcionalmente

(f) uno o más oligonucleótidos que consisten en una secuencia de nucleótidos seleccionada de las enumeradas en la Tabla 2 y Tabla 3.

6. El conjunto de sondas oligonucleotídicas de una cualquiera de las reivindicaciones 1 a 5, en el que uno o más de dichos oligonucleótidos está marcado con un resto para ayudar a la detección o manipulación, preferentemente en el que dicho resto es colorimétrico, quimioluminiscente, cromogénico, radiactivo, fluorescente, una enzima, un fragmento de anticuerpo, un marcador His, biotina o estreptavidina.

7. El conjunto de sondas oligonucleotídicas de una cualquiera de las reivindicaciones 1 a 6 en el que uno o más de dichos oligonucleótidos está inmovilizado en uno o más soportes sólidos, preferentemente marcados con un colorante o una pluralidad de colorantes, preferentemente un colorante luminiscente, en el que dicho soporte sólido se selecciona preferentemente de:

(i) partículas, láminas, geles, filtros, membranas, fibras, capilares, microplacas, tiras de microtitulación, portaobjetos, tubos, placas o pocillos; o

(ii) una partícula magnética, preferentemente una perla magnética.

8. Un método para realizar el perfil de la microbiota del tracto GI de un sujeto, comprendiendo dicho método:

(i) poner en contacto la muestra del tracto GI de dicho sujeto con un conjunto de sondas oligonucleotídicas como se define en una cualquiera de las reivindicaciones 1 a 7;

(ii) someter la muestra y el conjunto de sondas a condiciones que permiten la hibridación de las sondas con sus secuencias de nucleótidos diana dentro de moléculas de ácido nucleico en dicha muestra; y

(iii) para cada oligonucleótido en dicho conjunto de sondas, determinar la cantidad de su secuencia de nucleótidos diana que está presente en dicha muestra.

5 9. El método de la reivindicación 8 en el que cada oligonucleótido tiene un marcador unido al mismo, y la etapa (iii) comprende determinar, para cada oligonucleótido marcado, la cantidad de dicho marcador unido a dicha muestra determinando la fuerza de la señal del marcador que procede de la muestra.

10. El método de la reivindicación 8 en el que la etapa (iii) comprende

10 (a) marcar selectivamente las sondas oligonucleotídicas del conjunto de sondas cuando se hibridan con su secuencia de nucleótidos diana; y
(b) determinar la cantidad de cada sonda oligonucleotídica marcada producida en la etapa (a).

15 11. El método de la reivindicación 10, en el que se produce marcaje selectivo por extensión de cadena de la sonda oligonucleotídica con un nucleótido marcado, preferentemente un didesoxinucleótido marcado.

12. El método de la reivindicación 10, en el que dicho nucleótido marcado es un ddCTP marcado, preferentemente ddCTP marcado con biotina.

20 13. El método de una cualquiera de las reivindicaciones 10 a 12, en el que la etapa (b) comprende hibridación de los oligonucleótidos de la etapa de marcaje (a) a secuencias de nucleótidos complementarias de las sondas oligonucleotídicas.

25 14. El método de la reivindicación 13, en el que una o más de dichas secuencias de nucleótidos complementarias de las sondas oligonucleotídicas está inmovilizada en uno o más soportes sólidos, preferentemente marcados con un colorante de una pluralidad de colorantes, preferentemente un colorante luminiscente, en el que dicho soporte sólido se selecciona preferentemente de:

30 (i) partículas, láminas, geles, filtros, membranas, fibras, capilares, microplacas, tiras de microtitulación, portaobjetos, tubos, placas o pocillos; o
(ii) una partícula magnética, preferentemente una perla magnética.

15. El método de la reivindicación 8 en el que la etapa (iii) comprende

35 (a) realizar una reacción de amplificación de ácido nucleico dependiente de cebador; y
(b) para cada oligonucleótido en el conjunto de sondas, determinar la cantidad de producto de amplificación producido a partir del mismo en dicha reacción de amplificación de ácido nucleico dependiente de cebador.

40 16. Un método para preparar un perfil de microbiota convencional del tracto GI que es característico de una enfermedad o afección o un estadio de la misma, o el riesgo de desarrollar una enfermedad o afección, comprendiendo dicho método:

45 (i) identificar un sujeto con dicha enfermedad o afección o estadio de la misma o que esté en riesgo de desarrollar dicha enfermedad o afección y
(ii) realizar el método de una cualquiera de las reivindicaciones 8 a 15 en una muestra del tracto GI de dicho sujeto.

17. El método de una cualquiera de las reivindicaciones 8 a 16, en el que dicha muestra del tracto GI se selecciona de

50 (a) contenidos lumbinales del tracto GI, preferentemente contenidos del estómago, contenidos intestinales, moco y heces o combinaciones de los mismos,
(b) partes de la mucosa, la submucosa, la capa muscular externa, la adventicia y/o la serosa de un tejido / órgano del tracto GI,
(c) ácido nucleico preparado de (a) o (b), preferentemente por transcripción inversa y/o amplificación de ácido nucleico, y preferentemente en el que dicha muestra del tracto GI se obtiene del yeyuno, el íleon, el ciego, el colon, el recto o el ano.

18. Un método para diagnosticar o supervisar una enfermedad o afección en un sujeto o predecir o evaluar el riesgo de que un sujeto desarrolle una enfermedad o afección, comprendiendo dicho método:

60 (a) realizar el perfil de la microbiota del tracto GI de un sujeto con un método como se ha definido en una cualquiera de las reivindicaciones 8 a 17
(b1) comparar dicho perfil con un perfil de microbiota convencional del tracto GI que es característico de una enfermedad o afección o un estadio de la misma, o el riesgo de desarrollar una enfermedad o afección y/o
(b2) comparar dicho perfil con un perfil de microbiota anterior del tracto GI del sujeto; y
65 (c) determinar el grado de correlación entre dichos perfiles.

19. Un método para diagnosticar o supervisar una enfermedad o afección en un sujeto o predecir o evaluar el riesgo de que un sujeto desarrolle una enfermedad o afección, comprendiendo dicho método:

- 5 (a1) comparar los resultados de un método como se ha definido en una cualquiera de las reivindicaciones 8 a 17 con un perfil de microbiota convencional del tracto GI que es característico de una enfermedad o afección o un estadio de la misma o el riesgo de desarrollar una enfermedad o afección y/o
- (a2) comparar los resultados de un método como se ha definido en una cualquiera de las reivindicaciones 8 a 17 con un perfil de microbiota anterior del tracto GI del sujeto, y
- 10 (b) determinar el grado de correlación entre dichos perfiles, en el que el grado de correlación es indicativo de la presencia o ausencia de dicha enfermedad o afección, o el riesgo de desarrollar dicha enfermedad o afección, o el progreso de la enfermedad o afección.

20. El método de una cualquiera de las reivindicaciones 16 a 19, en el que dicha enfermedad o afección se selecciona de enfermedad inflamatoria del intestino, enfermedad de Crohn, colitis ulcerosa, síndrome del intestino irritable, cánceres del tracto GI, preferentemente cáncer de la boca, faringe, esófago, estómago, duodeno, yeyuno, íleon, ciego, colon, recto o ano, enfermedades atópicas, preferentemente eccema, asma, dermatitis atópica, conjuntivitis alérgica, rinitis alérgica y alergias alimentarias; trastornos metabólicos, preferentemente diabetes mellitus (tipo 1 y tipo 2), obesidad y síndrome metabólico; trastornos neurológicos, preferentemente depresión, esclerosis múltiple, demencia y enfermedad de Alzheimer; y autismo.

21. Un kit que comprende el conjunto de sondas oligonucleotídicas de las reivindicaciones 1 a 7, que comprende preferentemente además al menos uno de

- 25 (a) medio para marcaje selectivo de uno o más de los oligonucleótidos del conjunto de sondas,
- (b) soportes sólidos en los que uno o más de los oligonucleótidos del conjunto de sondas pueden inmovilizarse,
- (c) medio para realizar una amplificación y/o reacción de extensión de cebadores con uno o más de los oligonucleótidos del conjunto de sondas,
- (d) medio para realizar una reacción de transcripción inversa,
- (e) medio para tomar una muestra del tracto GI,
- 30 (f) medio para purificar o refinar una muestra de tracto GI,
- (g) medio para extraer ácido nucleico de una muestra del tracto GI,
- (h) medio para detectar ácido nucleico amplificado,
- (i) uno o más oligonucleótidos que hibridan selectivamente con ácidos nucleicos diana indicativos de cualquier otra enfermedad o afección médica asociada con la microbiota gastrointestinal, preferentemente enfermedad de Crohn,
- 35 enfermedad inflamatoria del intestino, colitis ulcerosa y síndrome del intestino irritable.

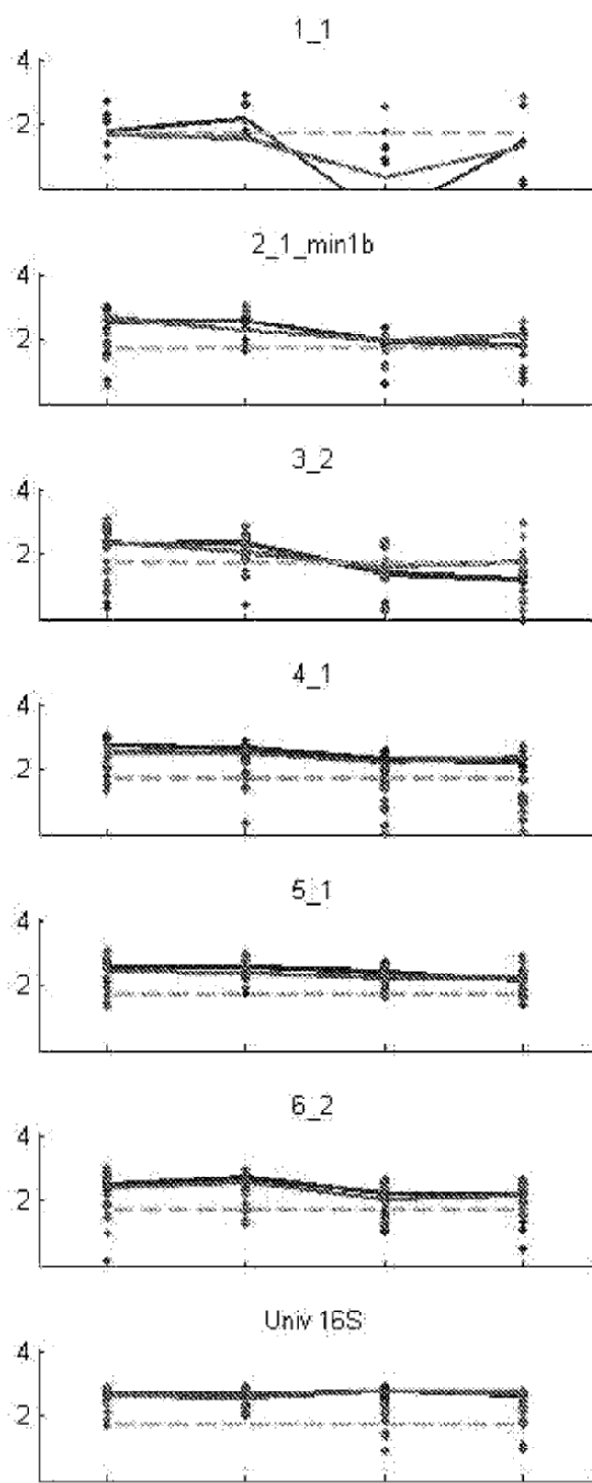


Figura 1

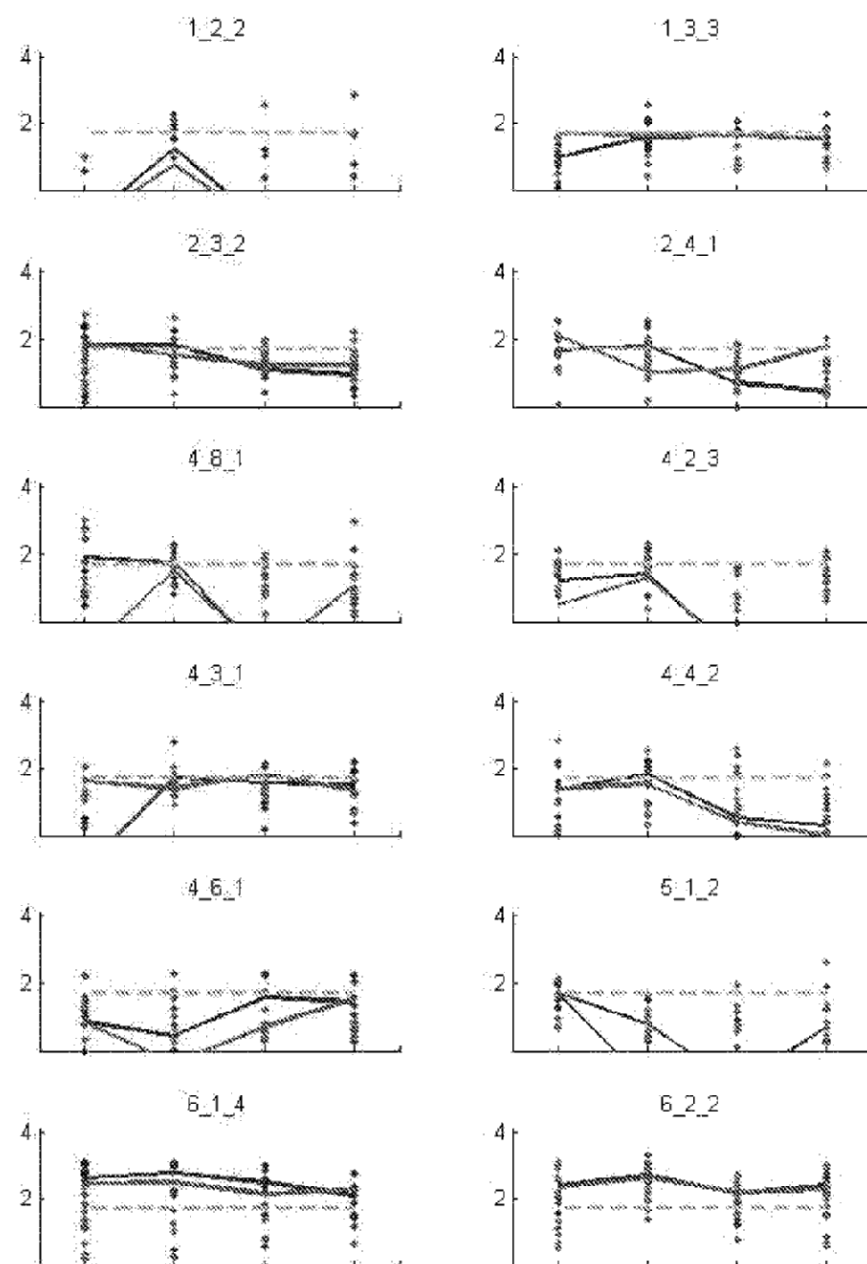


Figura 2