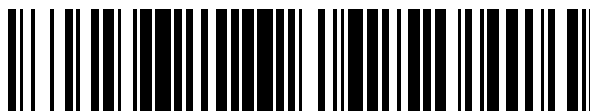


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 391 744**

51 Int. Cl.:  
**C12Q 1/68**

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **05824553 .1**

96 Fecha de presentación: **01.11.2005**

97 Número de publicación de la solicitud: **1815016**

97 Fecha de publicación de la solicitud: **08.08.2007**

54 Título: **Composiciones y procedimientos para el diagnóstico de trastornos del colon**

30 Prioridad:  
**01.11.2004 US 623771 P**  
**26.01.2005 US 646592 P**

45 Fecha de publicación de la mención BOPI:  
**29.11.2012**

45 Fecha de la publicación del folleto de la patente:  
**29.11.2012**

73 Titular/es:  
**GEORGE MASON UNIVERSITY (50.0%)**  
**Office of Technology Transfer, MSN-5G5, 4400**  
**University Drive**  
**Fairfax, VA 22030, US y**  
**RUSH UNIVERSITY (50.0%)**

72 Inventor/es:  
**GILLEVET, PATRICK**

74 Agente/Representante:  
**CURELL AGUILÁ, Mireia**

ES 2 391 744 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

Composiciones y procedimientos para el diagnóstico de trastornos del colon.

- 5 Esta solicitud reivindica los derechos de la solicitud provisional US nº 60/623.771, presentada el 1 de noviembre de 2004 y de la solicitud provisional US nº 60/646.592, presentada el 26 de enero de 2005.

### Antecedentes de la invención

- 10 La colitis ulcerosa y la enfermedad de Crohn son enfermedades inflamatorias crónicas del colon y del recto. Aunque los corticosteroides, los aminosalicilatos y los inmunomoduladores han proporcionado alguna ventaja en el tratamiento de la colitis ulcerosa, la anastomosis de la bolsa ileoanal por proctocolectomía restauradora (RP/IPAA) sigue siendo el tratamiento estándar para la gestión de la enfermedad esteroide-refractaria activa de forma prolongada. La complicación más corriente y debilitante de la IPAA es una inflamación sintomática de la bolsa del ileon o pouchitis. Estudios anteriores han demostrado una disminución significativa en la calidad de vida (IBDQ y SF-15  
15 36) cuando RP/IPAA se complica por pouchitis. La frecuencia de la pouchitis está entre el 30 y 50% hasta 5 años después de la operación, siendo la mayoría de los casos iniciales en los primeros 3-6 meses. Clínicamente, la pouchitis se caracteriza por aumento en la frecuencia de las deposiciones, urgencia fecal, hemorragia rectal y malestar general. Sin embargo, el diagnóstico de la pouchitis es una combinación de criterios clínicos, endoscópicos e histológicos. Existe un gran debate en cuanto a si la pouchitis es una extensión de la colitis ulcerosa o una enfermedad distinta. No ha habido ningún dato que favorezca fuertemente ninguna de las dos. Aunque se han propuesto muchas teorías, el mecanismo exacto de la enfermedad en pouchitis sigue siendo impreciso. La respuesta clínica espectacular a los antibióticos en la pouchitis sugiere que la microflora puede desempeñar una función causal. A pesar de una respuesta inicial a los antibióticos del 80%, el 60% de los pacientes presenta episodios de recaída de pouchitis y hasta el 30% de los pacientes desarrollan pouchitis crónica sintomática. No se han realizado estudios hasta la fecha para identificar ningún patrón u organismo específico de microflora en la patogenia de la pouchitis. En este estudio, los autores introducen Heterogeneidad de la Longitud del Amplicón, una nueva técnica de cultivo independiente para la caracterización detallada de la microflora en la pouchitis.
- 20 El documento WO03053220 describe un procedimiento de diagnóstico de la enfermedad de Crohn o la colitis ulcerosa utilizando una sonda que detecta la flagelina de *Helicobacter bilis* o una bacteria desconocida en la que un aumento indica un estado de la enfermedad.

- 35 Los cambios en la flora microbiana se han asociado la enfermedad inflamatoria intestinal y específicamente a la enfermedad de Crohn y la colitis ulcerosa (Seksik *et al.*, 2003, Gut 52:237-242; Ott *et al.*, 2004, Gut 53: 685-693; Kleesen *et al.*, *Scand. J. Gastroenterol.* 37: 1034-1041; Linskens *et al.*, 2002, *Scand. J. Gastroenterol.* 36 Suppl. 234: 29-40).

### Descripción de los dibujos

- 40 Figura 1. Histograma de tejido de Crohn y huellas con HLA del lumen.  
Figura 2. Histograma de tejido de colitis ulcerosa y huellas con HLA del lumen.  
Figura 3. Análisis principal de coordenadas (PCO) de huellas de Crohn y HLA de colitis ulcerosa.  
45 Figura 4. Histograma de bolsa normal y huellas con HLA de tejido de pouchitis.  
Figura 5. Análisis de coordenadas del principio (PCO) de huellas con HLA de pouchitis.  
Figura 6. Identificación de los picos en histograma de la bolsa normal y pouchitis.

### Descripción de la invención

- 50 La presente descripción se refiere a procedimientos y composiciones para el diagnóstico, seguimiento, pronóstico, análisis, etc., de enfermedades polimicrobianas. Una enfermedad polimicrobiana es una enfermedad o afección que está asociada a la presencia de al menos dos microorganismos diferentes, incluyendo, por ejemplo, las asociaciones entre bacteria-bacteria, virus-virus, parásito-parásito, bacteria-virus, bacteria-parásito y virus-parásito. Un procedimiento preferido de determinación de la población microbiana presente en una enfermedad polimicrobiana es la heterogeneidad de la longitud del amplicón ("HLA").

- 60 Los ejemplos de enfermedades polimicrobianas comprenden de manera no limitativa, por ejemplo, la coinfección de Borrelia y Ehrlichia en la borreliosis de Lyme; infecciones mixtas vírico-bacterianas durante epidemias de gripe; enfermedades respiratorias; gastroenteritis; conjuntivitis; queratitis, hepatitis, enfermedades periodontales, infecciones genitales; infecciones intra-abdominales; enfermedades intestinales inflamatorias; infecciones de las vías urinarias; infección necrosante de tejidos blandos.

- 65 La presente invención se refiere también a la población microbiana presente en el aparato digestivo y el lumen en pacientes normales, y en pacientes con enfermedades del aparato digestivo, especialmente enfermedades del colon, tales como la enfermedad inflamatoria intestinal, incluyendo la colitis ulcerosa, el síndrome de Crohn y la pouchitis. La presente invención se refiere especialmente a composiciones y procedimientos para diagnosticar,

pronosticar, y/o hacer el seguimiento de la evolución de la enfermedad de las enfermedades y afecciones mencionadas, por ejemplo, para determinar la presencia de la enfermedad en un paciente, para determinar un régimen terapéutico, para determinar la aparición de la enfermedad activa, para determinar la predisposición a la enfermedad, para determinar el curso de la enfermedad, etc.

La presente invención proporciona procedimientos para el diagnóstico y seguimiento de la evolución de la enfermedad de enfermedades inflamatorias intestinales, tales como la colitis ulcerosa, la enfermedad de Crohn o la pouchitis, que comprende determinar la presencia o ausencia de microbios, tales como bacterias, en una muestra de colon o lumen extraída de dicho paciente. La invención no se limita a la forma en que se lleva a cabo la determinación; se puede utilizar cualquier procedimiento adecuado. El término "microbio" incluye virus, bacterias, hongos y protistas. Aunque la descripción siguiente se puede redactar desde el punto de vista de bacterias, puede utilizarse cualquier microbio. La invención se refiere a un procedimiento para el diagnóstico de una enfermedad inflamatoria intestinal en un paciente, que comprende determinar el aumento o disminución en comparación con una muestra de mucosa normal de *Moraxella sp.* del grupo de las Pseudomonas, de *Comamonas sp.* del grupo Acidovorax, *Cryseobacterium sp.* del grupo Cytophaga, *Enterococcus sp.* del grupo de los Enterococcus, *Bacteroides sp.* del grupo Bacteroides, *Propionibacterium sp.* del grupo Propionibacterium, *Ruminococcus sp.* del grupo Clostridium Coccoides y *Bacteroides sp.* en una muestra de tejido de la mucosa del colon de dicho paciente, en el que

(a) la enfermedad inflamatoria intestinal es la enfermedad de Crohn y en el que un aumento de *Moraxella sp.*, *Comomonas sp.* y *Chryseobacter sp.* y una disminución de *Bacteroides sp.*, *Propionibacterium sp.*, *Enterococcus sp.* y *Ruminococcus sp.* indica la enfermedad de Crohn en dicho paciente, o en el que

(b) la enfermedad inflamatoria intestinal es la colitis ulcerosa y en el que un aumento de *Moraxella sp.*, *Comomonas sp.* y *Enterococcus sp.* y una disminución de *Bacteroides sp.*, *Propionibacterium sp.*, *Chryseobacter sp.* y *Ruminococcus sp.* indica colitis ulcerosa en dicho paciente.

El procedimiento de la invención puede comprender además determinar el aumento o disminución de una o más bacterias seleccionadas de entre el grupo constituido por *Bacteriodes sp.*, *Ruminococcus sp.*, *Moraxella sp.* y *Comomonas sp.* en una muestra del lumen del paciente.

La invención se refiere también a un procedimiento para diagnosticar pouchitis en un paciente de ileal-bolsa, que comprende detectar *Fusobacter sp.*, *Propionibacterium sp.* y al menos una de las bacterias del grupo de *E.coli* y/o *Shigella sp.* en una muestra de tejido de la mucosa del colon de dicho paciente, en el que un aumento de *Fusobacter sp.* y una disminución de una o más de las bacterias del grupo de *E. coli* y/o *Shigella sp.* en comparación con una muestra de bolsa normal y ausencia de *Propionibacterium sp.* es diagnóstico de pouchitis.

La presente descripción se refiere a cualquier composición o procedimiento que es adecuado para la detección de una población microbiana en una muestra (por ejemplo, de un paciente que padece una enfermedad polimicrobiana), tal como una muestra de aparato digestivo, lumen o heces. Una muestra de lumen es de interior del intestino.

Cualquier marcador que sea adecuado para identificar y distinguir un tipo microbiano puede utilizarse según la presente invención. Estos procedimientos pueden implicar la detección de ácido nucleico (por ejemplo, ADN, ARN, ARNm, ARNt, ARNr, etc.), proteína (por ejemplo, utilizando anticuerpos, reactivos de unión a proteínas), y cualquier otra biomolécula (por ejemplo, lípidos, hidratos de carbono, etc.) que sea útil para determinar específicamente la presencia o ausencia de bacterias en una muestra. Cualquier indicador variable o segmento no codificador (por ejemplo, elementos repetitivos, etc.) se pueden utilizar también, así como genes indicadores. Las regiones ITS puede utilizarse en los hongos.

También puede utilizarse métodos de cultivo convencionales, en los que se identifican bacterias y otros microorganismos cultivándoles en un medio, por ejemplo, utilizando un medio selectivo (por ejemplo, que comprende una fuente de carbono específica para bacterias) y/o en los que los microorganismos son identificados por sus características de crecimiento, morfología y otros criterios aliados típicos utilizados para determinar la identidad celular y clasificación filogenética. Cualquiera de estos métodos también se pueden utilizar en combinación con métodos citológicos e histológicos, con los que se pueden teñir y observar muestras de biopsia o muestras cultivadas (por ejemplo, por corte, o por montaje en un portaobjetos u otro portador).

Como se ha mencionado, las composiciones y procedimientos son útiles para diagnóstico y pronóstico relacionados con enfermedades polimicrobianas, tales como enfermedades inflamatorias del intestino, incluyendo la colitis ulcerosa, la enfermedad de Crohn y la pouchitis. Se pueden utilizar marcadores y huellas para diagnosticar las enfermedades, y distinguir las de otras enfermedades del aparato digestivo. También se puede utilizar para evaluar el estado de la enfermedad, la gravedad y el pronóstico, solos, o en combinación con otras pruebas. Por ejemplo, los marcadores se pueden utilizar en combinación con el índice de actividad de la enfermedad de Crohn (CAI) o los criterios de Trulove y Witts para evaluar la actividad de la enfermedad en la colitis ulcerosa. La información sobre el estado de los microorganismos se puede utilizar también para determinar cuándo iniciar el tratamiento farmacéutico u otros regímenes terapéuticos.

- Los procedimientos y composiciones también se pueden utilizar para controlar el curso de la enfermedad en un paciente bajo tratamiento o para seguir la evolución de la enfermedad, independientemente del régimen de tratamiento. Por ejemplo, los pacientes con síndromes inflamatorios del intestino pueden presentar remisiones espontáneas o provocadas por fármacos. Para seguir el curso de la remisión y determinar cuando la enfermedad está activa, se pueden tomarse muestras periódicamente, y analizarse para determinar el aspecto de los marcadores microbianos específicos o huellas en el tejido intestinal, lumen, lavado de colon, muestras de mucosa o heces.
- La evaluación de la población microbiana se puede realizar en cualquier muestra extraída de un paciente, incluyendo de lumen, lavado del colon, tejido intestinal, mucosa intestinal, tejido gástrico, mucosa gástrica, heces, etc. Pueden extraerse muestras de cualquier parte del aparato digestivo, especialmente de los intestinos delgado y grueso. El intestino grueso o colon es la parte del intestino desde el ciego hasta el recto. Se divide en ocho secciones: el ciego, el apéndice, el colon ascendente, el colon transverso, el colon descendente, el colon sigmoideo, el recto y el ano. Un lavado de colon es el líquido que queda en el intestino después de administrar a un paciente un laxante. La mucosa intestinal es el revestimiento en superficie del aparato intestinal. Los sujetos incluyen, por ejemplo, animales, seres humanos, primates no humanos, mamíferos, monos, ganado, ovejas, cabras, cerdos, animales domésticos (por ejemplo, perros, gatos), animales pequeños, reptiles, aves, etc.
- Cualquier procedimiento adecuado puede utilizarse para extraer muestras del intestino. La biopsia endoscópica es un procedimiento frecuente en la que se inserta un endoscopio de fibra óptica en el aparato digestivo a través de un orificio natural del cuerpo. Se observa directamente el revestimiento del intestino y se pellizca una muestra con fórceps acoplados a un cable largo que discurre en el interior del endoscopio. Los endoscopios e instrumentos adecuados para la extracción de muestras de biopsia son bien conocidos, e incluyen los dados a conocer en, por ejemplo, las patentes US nº 6.632.182 y nº 6.443.909.

La Tabla 3 resume las bacterias que se han detectado en tejido de mucosa y el lumen de los pacientes de referencia, y los pacientes que tienen la enfermedad de Crohn o colitis ulcerosa. La Tabla 5 resume las bacterias que se han detectado en la mucosa y el lumen de pacientes que tienen pouchitis y pouchitis de referencia (pacientes con proctocolectomía restauradora, pero sin complicaciones posoperatorias). Se clonaron amplicones de RCP y se secuenciaron de estas muestras. En resumen, se extrajo ADN de cada muestra mezclada. La mezcla de ADN de la mucosa comprendía ADN de las células de la mucosa gastrointestinal y otras células gastrointestinales, así como de las bacterias. Las primeras dos regiones variables del ARN ribosómico 16S se ampliaron utilizando cebadores universales eubacterianos. Posteriormente, la mezcla de ampliación se separó y se caracterizó en un gel de toma de huellas. La imagen resultante del gel o tabular la compilación de los datos (véanse, por ejemplo, las Tablas 1, 2 y 4) - que comprenden bandas individuales, discretas (amplicones de RCP) - puede denominarse la "huella de la HLA". La huella digital de la HLA puede caracterizarse además identificando la longitud de cada uno de los replicones que la componen y/o sus secuencias de nucleótidos específicas. Los amplicones de la población microbiana pueden clonarse y secuenciarse a continuación, cuando la secuencia se correlaciona con un grupo específico de bacterias, especies o cepa. Utilizando este procedimiento, la abundancia de los clones de cada especie es proporcional a su abundancia en la población correspondiente, y se puede correlacionar con los picos de la huella digital de la HLA. Los datos de la secuencia se pueden utilizar para buscar en la base de datos ribosómica (RDP) utilizando una herramienta de búsqueda de secuencias convencional (Megablast) disponible en el Centro Nacional de Información Biotecnológica (NCBI) en el NIH. Véase, por ejemplo, J.R. Cole, Chai B., Marsh T.L., Farris R.J., Wang Q., Kulam S.A., Chandra S., McGarrell D.M., Schmidt T.M., Garrity G.M., Tiedje J.M.. El proyecto de base de datos ribosómica (RDP-II): anticipo de un nuevo autoalineador que permite actualizaciones periódicas y la nueva taxonomía procariota. *Nucleic Acids Res.* 1 de enero de 2003, 31(1):442-3. El número RDP obtenido de los resultados de la búsqueda se puede analizar utilizando una secuencia PERL habitual para clasificar la División, Subdivisión, Grupo y Subgrupo de cada clon, y los resultados pueden tabularse, e importarse en Excel u otras bases de datos adecuadas.

#### 50 Procedimientos de detección de ácidos nucleicos

- Los procedimientos de detección tienen una variedad de aplicaciones, incluyendo aplicaciones para diagnóstico, pronóstico, forense y para investigación. Para llevar a cabo la detección de ácidos nucleicos, un polinucleótido según la presente invención se puede utilizar como "sonda". El término "sonda" o "sonda de polinucleótido" tiene su significado habitual en la técnica, por ejemplo, un polinucleótido que es eficaz para identificar (por ejemplo, por hibridación), cuando se utiliza en un proceso apropiado, la presencia de un polinucleótido diana para el que está diseñado. La identificación puede implicar simplemente determinar la presencia o ausencia, o puede ser cuantitativo, por ejemplo, en la evaluación de las cantidades de un polinucleótido (por ejemplo, copias de un ARN ribosómico) presentes en una muestra. Como se explica con más detalle más adelante, se puede utilizar cualquier procedimiento adecuado, incluyendo, pero sin limitarse a, HLA, RCP, secuenciación de nucleótidos, inmunotransferencia Southern y o micromatrices de ADN (por ejemplo, cuando una micromatriz comprende un gran número de secuencias específicas para una o más bacterias de la presente invención).
- Pueden utilizarse ensayos que permiten la cuantificación y/o detección de la presencia o ausencia de un ácido nucleico diana en una muestra. Los ensayos pueden llevarse a cabo a nivel de una sola célula, o en una muestra

que comprende muchas células, en donde el ensayo es la expresión "promedio" de toda el conjunto completo de células y tejido presente en la muestra. Se puede utilizar cualquier formato de ensayo adecuado, incluyendo, pero sin limitarse a, por ejemplo, análisis de inmunotransferencia Southern, análisis de inmunotransferencia Northern, reacción en cadena de la polimerasa ("RCP") (por ejemplo, Saiki *et al.*, *Science*, 241:53, 1988; patente US nº 4.683.195, nº 4.683.202 y nº 6.040.166; *PCR Protocols: A guide to Methods and Applications*, Innis *et al.*, eds, Academic Press, Nueva York, 1990), reacción en cadena de la polimerasa con transcriptasa inversa ("RT-RCP"), RCP anclada, ampliación rápida de extremos del ADNc ("RACE") (por ejemplo, Schaefer en *Gene Cloning and Analysis: Current Innovations*, páginas 99-115, 1997), reacción en cadena de la ligasa ("RCL") (patente EP 320 308), RCP por una sola cara (Ohara *et al.*, *Proc. Natl. Acad. Sci.*, 86:5673-5677, 1989), procedimientos de indexación (por ejemplo, patente US nº 5.508.169), hibridación *in situ*, presentación diferencial (por ejemplo, Liang *et al.*, *Nucl. Acid. Res.*, 21:3269-3275, 1993; patente US nº 5.262.311, nº 5.599.672 y nº 5.965.409; documento WO 97/18454; Prashar y Weissman, *Proc. Natl. Acad. Sci.*, 93:659-663, y patente US nº 6.010.850 y nº 5.712.126; Welsh *et al.*, *Nucleic Acid Res.*, 20:4965-4970, 1992, y patente US nº 5.487.985) y otras técnicas de huellas de ARN, ampliación basada en la secuencia de ácido nucleico ("NASBA") y otros sistemas de ampliación basados en la transcripción (por ejemplo, patente US nº 5.409.818 y nº 5.554.527; documento WO 88/10315), matrices de polinucleótidos (por ejemplo, patente US nº 5.143.854, nº 5.424.186; nº 5.700.637, nº 5.874.219 y nº 6,054,270; documentos PCT WO 92/10092; PCT WO 90/15070), Qbeta Replicasa (documento PCT/US87/00880), Strand Ampliación con Desplazamiento de cadena ("SDA"), Reacción en Cadena de reparación ("RCR"), ensayos de protección de nucleasa, los procedimientos basados en sustracción, Rapid-Scan™, etc. Otros procedimientos útiles comprenden de manera no limitativa, por ejemplo, procedimientos de ampliación con plantillas, RCP competitiva (por ejemplo, patente US nº 5.747.251), ensayos basados en redox (por ejemplo, patente US nº 5.871.918), Ensayos basados en Taqman (por ejemplo, Holland *et al.*, *Proc. Natl. Acad. Sci.*, 88:7276-7280, 1991; patente US nº 5.210.015 y nº 5.994.063), seguimiento por fluorescencia en tiempo real (por ejemplo, patente US nº 5.928.907), etiquetas moleculares de transferencia de energía (por ejemplo, patente US nº 5.348.853, nº 5.532.129, nº 5.565.322, nº 6.030.787 y nº 6.117.635; Tyagi y Kramer, *Nature Biotech.*, 14:303-309, 1996). Cualquier procedimiento adecuado para el análisis de células individuales de expresión de polinucleótido o de proteínas se pueden utilizar, incluyendo la hibridación *in situ*, inmunocitoquímica, MACS, FACS, citometría de flujo, etc. Para los ensayos de una sola célula, los productos de expresión se pueden medir utilizando anticuerpos, RCP u otros tipos de ampliación de ácido nucleico (por ejemplo, Brady *et al.*, *Methods Mol. & Cell. Biol.* 2, 17-25, 1990; Eberwine *et al.*, 1992, *Proc. Natl. Acad. Sci.*, 89, 3010-3014, 1992; patente US nº 5.723.290). Estos y otros métodos se puede llevar a cabo de manera convencional, por ejemplo, como se describe en las publicaciones mencionadas.

Muchos de dichos procedimientos pueden requerir que el polinucleótido esté marcado, o comprenda un tipo de nucleótido concreto útil para la detección. La presente invención incluye dichos polinucleótidos modificados que son necesarios para llevar a cabo dichos procedimientos. Por lo tanto, los polinucleótidos pueden ser ADN, ARN, híbridos ADN:ARN, APN, etc., y pueden comprender cualquier modificación o sustituyente que sea eficaz para conseguir la detección.

La presente invención proporciona procedimientos para diagnosticar o pronosticar la colitis ulcerosa, la enfermedad de Crohn o la pouchitis, en un paciente, que comprende, uno o más de las siguientes etapas en cualquier orden eficaz, por ejemplo, poner en contacto un tejido gastrointestinal o una muestra de lumen que comprende ácido nucleico con una sonda de polinucleótidos que es específica para al menos una bacteria en condiciones eficaces para que dicha sonda se hibride específicamente con dicho ácido nucleico, y detectar la hibridación entre dicha sonda y dicho ácido nucleico, en los que la presencia de una o más bacterias seleccionadas de entre el grupo siguiente dicha bacteria indica la presencia de la enfermedad o el estado de enfermedad de la colitis ulcerosa, la enfermedad de Crohn o la pouchitis. El procedimiento puede comprender además la obtención de una muestra de colon, por ejemplo, por biopsia endoscópica y/o la extracción del ácido nucleico de la muestra.

La frase "específico para" un microbio tiene un significado funcional que indica que la sonda (o anticuerpo si se utiliza en un contexto de proteínas) se puede utilizar para identificar la presencia del microbio objetivo en una muestra y distinguirlo del microbio que no es objetivo. También es específico en el sentido de que puede utilizarse para detectar microbios objetivo por encima del ruido de fondo ("unión no específica"). Esta misma definición es aplicable también a una sonda de polinucleótido o anticuerpo. Las sondas también pueden ser descritas como específicas para una secuencia, donde una secuencia específica es un orden definido de nucleótidos (o secuencias de aminoácidos, si se trata de una secuencia de polipéptido) que se produce en el polinucleótido.

La frase "hibridar específicamente" indica que la hibridación entre polinucleótidos monocatenarios se basa en la complementariedad de secuencia de nucleótidos. Las condiciones eficaces se seleccionan de tal manera que la sonda se hibrida a un ácido nucleico diana preseleccionada y/o definida en la muestra. Por ejemplo, si se desea la detección de un polinucleótido de un ARN ribosómico, se puede seleccionar una sonda que puede hibridarse con el ARN ribosómico diana en condiciones de alta severidad, sin hibridación significativa con otras secuencias que no son diana en la muestra. Por ejemplo, pueden seleccionarse de forma rutinaria las condiciones que requieren 100% o complementariedad completa entre la diana y la sonda.

La puesta en contacto de la muestra con la sonda puede llevarse a cabo por cualesquier medios eficaces en cualesquier medios eficaces. Se puede llevar a cabo en una matriz sólida, líquida, congelada, gaseosa, amorfa,

solidificada, coagulada, coloide, etc., mezclas de las mismas. Por ejemplo, una sonda en un medio acuoso puede ponerse en contacto con una muestra que está también en un medio acuoso, o que está fijada a una matriz sólida, o viceversa.

5 Generalmente, tal como se utiliza en toda la memoria, la expresión "condiciones eficaces" significa, por ejemplo, el medio concreto en el que se logra el efecto deseado, tal como la hibridación entre una sonda y su diana, o la unión del anticuerpo a una proteína diana. Dicho medio incluye, por ejemplo, tampones adecuados, agentes oxidantes, agentes reductores, pH, cofactores, temperatura, concentraciones de iones, edad adecuada y/o etapa de la célula (tales como, en particular la parte del ciclo celular, o en una etapa específica en la que se están expresando los genes específicos) en la que las células se están utilizando, condiciones de cultivo (incluyendo substrato, oxígeno y dióxido de carbono, etc.) Cuando la hibridación es el medio elegido para conseguir la detección, la sonda y la muestra se pueden combinar de tal manera que las condiciones resultantes son funcionales para que dicha sonda se hibride específicamente al ácido nucleico en dicha muestra.

15 Para detectar la presencia de una sonda específicamente hibridada a una diana, se puede utilizar cualquier procedimiento adecuado. Por ejemplo, los polinucleótidos se pueden marcar utilizando marcadores radiactivos tales como  $^{32}\text{P}$ ,  $^{35}\text{S}$ ,  $^3\text{H}$  o  $^{14}\text{C}$ , por mencionar algunos trazadores utilizados frecuentemente. se puede utilizar también marcaje no radiactivo, por ejemplo, biotina, avidina, digoxigenina, antígenos, enzimas o sustancias que tienen propiedades físicas detectables, tales como la fluorescencia o la emisión o absorción de luz a una longitud de onda deseada, etc.

25 Se puede utilizar cualquier muestra de ensayo en la que se desea identificar la presencia o ausencia de bacterias, incluyendo, por ejemplo, sangre, orina, saliva, lumen (para la extracción de ácido nucleico, véase, por ejemplo, patente US nº 6.177.251), heces, extensiones compuestas de tejido, tejido de biopsia, secciones de tejidos, células cultivadas, lavado intestinal, lavado del colon, mucosa intestinal, etc.

30 Los resultados para cualquiera de los ensayos mencionados en la presente memoria (incluyendo los ensayos en otras secciones a continuación) pueden ser con respecto a una muestra de referencia. Por ejemplo, un aumento o disminución puede ser con respecto (en comparación) a un lumen normal o muestra de mucosa. La muestra normal puede ser del mismo paciente, pero de una región o período no aquejado (por ejemplo, cuando el paciente está en remisión). También puede ser de un valor estándar que se calcula basándose en una población normalizada de individuos. Pueden utilizarse estadísticas estándar para determinar si los valores son significativos.

35 La presente descripción también proporciona procedimientos para la toma de huellas de ácido nucleico de la población de microbios presentes en una muestra, por ejemplo, utilizando cebadores universales para los microorganismos en cuestión, si son eubacterias, arqueobacterias, hongos o protistas. Puesto que cada muestra contiene una población de microbios característica que es representativa de la enfermedad, el muestreo de los ácidos nucleicos de los microbios puede producir una variedad característica de fragmentos de polinucleótidos asociados a la enfermedad. Estos pueden presentarse con cualquier característica física, incluyendo el tamaño, la secuencia, la movilidad, el peso molecular (por ejemplo, utilizando espectroscopia de masas), etc. Cualquier procedimiento de toma de huellas digitales pueden utilizarse, incluyendo, por ejemplo, AFLP, HLA, LH-RCP, ARISA, RAPD, etc. Las tablas 1, 2 y 4 muestran la frecuencia de los amplicones en diversas muestras de referencia y enfermedad. Aunque un amplición específico puede no ser diagnóstico de la afección el 100% de las veces, la utilización de amplicones múltiples aumenta la seguridad diagnóstica. Por otra parte, cuando se está haciendo el seguimiento de una afección, puede ser ventajoso hacer el seguimiento de una huella digital compleja (tal como se muestra en la Tabla 1) a medida que se diferencia de un tiempo de muestreo a otro.

45 A lo largo de estas líneas, la presente descripción proporciona un procedimiento para diagnosticar, pronosticar, o hacer el seguimiento de la evolución de una enfermedad polimicrobiana (por ejemplo, una enfermedad inflamatoria del intestino, como la colitis ulcerosa, pouchitis, o enfermedad de Crohn), que comprende uno o más de las siguientes etapas en cualquier orden eficaz, por ejemplo, realizando una reacción de ampliación en una muestra que comprende ácido nucleico con al menos dos cebadores sonda de polinucleótidos que son eficaces para la ampliación de la población microbiana presente en dicha muestra, y detectando los productos de reacción de dicha reacción de ampliación, por lo que dichos productos de reacción comprenden un patrón que indica la presencia de la enfermedad o el estado de la enfermedad.

50 Por "estado de enfermedad", se entiende la afección relativa de la enfermedad en comparación con su estado en un momento anterior. Por ejemplo, cuando los productos de reacción de la muestra son diferentes (por ejemplo, en cantidad o tamaño) de un período de gravedad de la enfermedad, esto indicaría que el estado de la enfermedad del paciente había cambiado. Los productos de reacción pueden presentar una diferencia antes de que el paciente realmente manifieste síntomas de la enfermedad, y por lo tanto se puede utilizar como pronóstico para predecir una recaída. Asimismo, un cambio en los productos de reacción también puede indicar que la enfermedad está mejorando y/o responder a un régimen de tratamiento.

65 El término "ampliación" indica que las secuencias de ácido nucleico aumentan en número de copias a una suma o cantidad a la que se pueden detectar. La ampliación se puede llevar a cabo convencionalmente, utilizando cualquier

técnica adecuada, incluyendo la reacción en cadena de la polimerasa (RCP), NASBA (por ejemplo, utilizando T7 ARN polimerasa), RCL (reacción en cadena de ligadura), LH-RCP, ARISA.

Los ácidos nucleicos totales se puede extraer de una muestra, o la muestra puede ser tratada de tal manera que se extraiga preferentemente ácido nucleico sólo de los microbios que están presentes en ella. Las extracciones de ADN se pueden realizar con los kits disponibles en el mercado, tales como el kit Bio101 de Qbiogene, Inc., Montreal, Quebec. Para evitar la contaminación por múltiples muestras durante el proceso de homogeneización de una muestra, cada muestra se puede procesar por separado y completamente dando lugar a extracciones de ADN con alto rendimiento.

En determinadas formas de realización de la presente invención, se puede utilizar ARN ribosómico ("ARNr") para distinguir y detectar bacterias. Por ejemplo, los ribosomas bacterianos se componen de una subunidad pequeña y una grande, cada una de las cuales está compuesto además de ARN ribosómicos y proteínas. El ARNr de la subunidad pequeña puede denominarse SSU ARNr, y de la subunidad mayor LSU ARNr. Un gran número de ARNr se han secuenciado, y estos están disponibles al público en varias bases de datos accesibles. Véase, por ejemplo, Wuyts *et al.*, The European database on small subunit ribosomal RNA, *Nucleic Acids Res.*, 30, 183-185, 2002; Cole *et al.*, The Ribosomal Database Project (RDP-II): previewing a new autoaligner that allows regular updates and the new prokaryotic taxonomy. *Nucleic Acids Res.*, 31 (1): 442-3, 2003. Véase, también, <http://rdp.cme.msu.edu/html/> consultado el 14 de junio de 2004. Cualquier ARNr puede utilizarse como marcador, incluyendo, pero sin limitarse a, 16S, 23S y 5S.

Las secuencias de cebador para ARNr pueden diseñarse de forma rutinaria para detectar especies específicas de bacterias, o para detectar grupos de bacterias, por ejemplo, cuando una secuencia conservada es característica de un grupo de bacterias. HLA-RCP puede llevarse a cabo de forma rutinaria, por ejemplo, utilizando un cebador directo marcado por fluorescencia 27F (5'-[6FAM] AGAGTTTGATCCTGGCTCAG-3') (SEQ ID n° 37) y cebador inverso sin marcar 338R' (5'-GCTGCCTCCCGTAGGAGT-3') (SEQ ID n° 38). Ambos cebadores son muy específicos para las eubacterias (Lane, D.J., 16S/23S rRNA Sequencing, in *Nucleic Acid Techniques in Bacterial Systematics*, E.S.a.M. Goodfellow, Editor. 1991, John Wiley & Sons Ltd: West Sussex, Inglaterra. págs. 115-175).

Pueden utilizarse también cebadores que amplían la región correspondiente en Archae (Burggraf, S., T. Mayer, R. Amann, S. Schadhauer, C.R. Woese y K.O. Stetter, Identifying Members of the Domain Archaea with rRNA-Targeted Oligonucleotide Probes. *App. Environ. Microbiol.*, 1994. 60: págs. 3112-3119), Eucariotas (Rowan, R. y D.A. Powers, Ribosomal RNA sequences and the diversity of symbiotic dinoflagellates (zooxanthellae). *Proceedings of the National Academy of Sciences, USA*, 1992. 89: págs. 3639-3643) y hongos (Borneman, J. y J. Hartin, PCR primers that amplify fungal rRNA genes from Environmental Samples. *App. Environ. Microbiol.*, 2000. 66 (10): págs. 4356-4360).

Pueden seleccionarse cebadores de cualquier ácido nucleico del agente infeccioso, incluyendo de ARNr, ARNt, ADN genómico, etc. Los cebadores pueden ser para regiones variables, hélices, regiones conservadas, etc.

Los cebadores seleccionados pueden utilizarse en la heterogeneidad de la longitud del amplicón ("HLA") para generar huellas digitales que caracterizan a la población bacteriana (Ritchie, N.J., *et al.*, Use of Length Heterogeneity PCR and Fatty Acid Methyl Ester Profiles to Characterize Microbial Communities in Soil. *Applied and Environmental Microbiology*, 2000. 66(4): págs. 1668-1675; Suzuki, M. M.S. Rappe y S.J. Giovannoni, Kinetic bias in estimates of coastal picoplankton community structure obtained by measurements of small-subunit rRNA gene PCR amplicon length heterogeneity. *Applied and Environmental Microbiology* [*Appl. Environ. Microbiol.*]. 1998. 64(11): págs. 4522-4529; Litchfield, C.D. y P.M. Gillevet, Microbial diversity and complexity in hypersaline environments: A preliminary assessment. *Journal of Industrial Microbiology & Biotechnology* [*J. Ind. Microbiol. Biotechnol.*]. 2002. 28(1): págs. 48-55; Mills, D.K., *et al.*, A Comparison of DNA Profiling Techniques for Monitoring Nutrient Impact on Microbial Community Composition during Bioremediation of Petroleum Contaminated Soils. *J. Microbiol. Method*, 2003. 54: págs. 57-74).

Pueden utilizarse cebadores individuales o una mezcla, por ejemplo, que comprende secuencias degeneradas, secuencias de uno o más grupos, reacción múltiple en la que se evalúan diferentes grupos utilizando cebadores marcados con diferentes etiquetas fluorescentes, etc.

Los productos de reacción (es decir, los fragmentos que se detectan después de la reacción de ampliación) pueden analizarse por análisis estadístico, como por ejemplo el análisis PCO (véase Ejemplos) para determinar qué productos que son diagnóstico de la enfermedad.

#### Detección de polipéptidos

La presente descripción también proporciona composiciones y métodos para detectar polipéptidos y otras biomoléculas que son característicos de la población microbiana. Por ejemplo, la presente descripción proporciona procedimientos para diagnosticar o pronosticar la colitis ulcerosa, pouchitis o enfermedad de Crohn, que comprende: una o más de las etapas siguientes en cualquier orden eficaz, por ejemplo, poner en contacto

una muestra que comprende proteína con un anticuerpo que es específico para una bacteria en condiciones eficaces para que dicho anticuerpo se una específicamente a dichas bacterias, y detectar la unión entre dicho anticuerpo y dicha bacteria.

5 Los polipéptidos se pueden detectar, observar, determinar, cuantificar, etc. según cualquier método eficaz. Los métodos útiles comprenden, por ejemplo, de manera no limitativa, inmunoanálisis, RIA (radioinmunoanálisis), ELISA, (ensayo inmunosorbente con enzima ligada), inmunofluorescencia, citometría de flujo, histología, microscopía electrónica, microscopía de luz, análisis *in situ*, inmunoprecipitación, inmunotransferencia Western, etc.

10 Pueden realizarse inmunoanálisis en líquidos o sobre soporte biológico. Por ejemplo, una muestra (por ejemplo, sangre, lumen, orina, células, tejido, fluido del bulbo raquídeo, fluidos corporales, etc.) se puede poner en contacto con e inmovilizarse sobre un soporte en fase sólida o un vehículo tal como nitrocelulosa, u otro sólido compatible que es capaz de inmovilizar células, partículas celulares o proteínas solubles. El soporte se puede lavar después con tampones adecuados seguido de tratamiento con el anticuerpo específico de la bacteria marcada de forma  
15 detectable. El soporte en fase sólida puede lavarse a continuación con un tampón una segunda vez para eliminar el anticuerpo no unido. La cantidad de marcador unido sobre un soporte sólido se puede detectar entonces por medios convencionales.

20 Un "soporte o portador en fase sólida" incluye cualquier soporte capaz de unirse a un antígeno, anticuerpo u otro acompañante de unión específica. Los soportes o portadores incluyen vidrio, poliestireno, polipropileno, polietileno, dextrano, nilón, amilasas, celulosas naturales y modificadas, poliácridamidas, y magnetita. Un material de soporte puede tener cualquier configuración estructural o física. Por lo tanto, la configuración del soporte puede ser esférica, como en una perla, o cilíndrica, como en la superficie interior de un tubo de ensayo, o la superficie externa de una varilla. Alternativamente, la superficie puede ser plana tal como una lámina, tira de ensayo, etc. Los soportes  
25 preferidos incluyen perlas de poliestireno.

Una de las muchas maneras en que un anticuerpo específico de bacterias puede marcarse de forma detectable es enlazándolo a una enzima y utilizándolo en un inmunoanálisis enzimático (EIA). Véase, por ejemplo, Voller, A., "The Enzyme Linked Immunosorbent Assay (ELISA)", 1978, *Diagnostic Horizons* 2, 1-7, Microbiological Associates Quarterly Publication, Walkersville, Md.); Voller, A. *et al.*, 1978, *J. Clin. Pathol.* 31, 507-520; Butler, J. E., 1981, *Meth. Enzymol.* 73,482-523; Maggio, E. (ed.), 1980, *Enzyme Immunoassay*, CRC Press, Boca Raton, Fla.. La enzima que se une al anticuerpo reaccionará con un sustrato apropiado, preferentemente un sustrato cromógeno, de tal manera como para producir un resto químico que puede detectarse, por ejemplo, por medios espectrofotométricos fluorimétricos o visualmente. Las enzimas que se pueden utilizar para marcar de forma detectable el anticuerpo  
30 comprenden de manera no limitativa, malato deshidrogenasa, estafilococo nucleasa, delta-5-esteroide isomerasa, alcohol de levadura deshidrogenasa, .alfa.-glicerofosfato, deshidrogenasa, triosa fosfato isomerasa, peroxidasa de rábano picante, fosfatasa alcalina, asparaginasa glucosa oxidasa, .beta.-galactosidasa, ribonucleasa, ureasa, catalasa, glucosa-6-fosfato deshidrogenasa, glucoamilasa y acetilcolinesterasa. La detección puede realizarse por métodos colorimétricos que emplean un sustrato cromógeno para la enzima. La detección también puede realizarse  
40 por comparación visual del grado de reacción enzimática de un sustrato en comparación con estándares preparados de manera similar.

La detección también puede realizarse utilizando cualquiera de varios otros inmunoanálisis. Por ejemplo, mediante marcado radiactivo de los anticuerpos o fragmentos de anticuerpos, es posible los péptidos mediante la utilización  
45 de un radioinmunoanálisis (RIA). Véase, por ejemplo, Weintraub, B., *Principles of Radioimmunoassays*, Seventh Training Course on Radioligand Assay Techniques, The Endocrine Society, Marzo, 1986. El isótopo radioactivo puede detectarse por medios tales como la utilización de un contador gamma o un contador de centelleo o por autorradiografía.

50 También es posible marcar el anticuerpo con un compuesto fluorescente. Cuando el anticuerpo marcado con fluorescencia se expone a luz de la longitud de onda apropiada, su presencia puede detectarse entonces debido a la fluorescencia. Entre los compuestos marcadores fluorescentes más comúnmente utilizados están el isotiocianato de fluoresceína, rodamina, ficoeritrina, ficocianina, aloficocianina, o-ftaldehído y fluorescamina. El anticuerpo también puede marcarse de forma detectable utilizando metales emisores de fluorescencia tales como los de la serie de los  
55 lantánidos. Estos metales pueden unirse al anticuerpo utilizando grupos quelantes metálicos tales como ácido dietilentriamin-pentaacético (DTPA) o ácido etilendiamintetraacético (EDTA).

El anticuerpo también puede marcarse de manera detectable por acoplamiento a un compuesto quimioluminiscente. La presencia del anticuerpo marcado con quimioluminiscencia se determina entonces detectando la presencia de  
60 luminiscencia que aparece durante el curso de una reacción química. Ejemplos de compuestos marcadores quimioluminiscentes útiles son el luminol, isoluminol, éster de acridinio teromático, imidazol, sal de acridinio y éster de oxalato.

Del mismo modo, un compuesto bioluminiscente puede utilizarse para marcar el anticuerpo de la presente invención.  
65 La bioluminiscencia es un tipo de quimioluminiscencia que se encuentra en sistemas biológicos en los que una proteína catalítica aumenta la eficacia de la reacción quimioluminiscente. La presencia de una proteína



bioluminiscente se determina detectando la presencia de luminiscencia. Los compuestos bioluminiscentes importantes para etiquetado son luciferina, luciferasa y acquirin.

La presente descripción también se refiere a la prevención y/o tratamiento de afecciones inflamatorias del intestino, en un paciente que lo necesita, que comprende la administración de antibióticos, así como otros compuestos antibacterianos, que son producidos por las bacterias en el aparato digestivo de individuos normales. Puede utilizarse un procedimiento probiótico, en el que se administran las bacterias que producen estos compuestos, en vez de proporcionar los compuestos en formas purificadas.

Como se ha descrito con detalle anteriormente, la población microbiana de pacientes con afecciones inflamatorias del intestino es perturbada. Estas perturbaciones pueden tener profundas consecuencias en la salud del paciente. Algunas bacterias, tales como *Ruminococcus sp.* producen antibióticos que tienen efectos protectores y antibacterianos sobre las bacterias patógenas. Por ejemplo, se muestra anteriormente en la Tabla 5 anterior que *R. gnavus* se reduce en pacientes que padecen la enfermedad de Crohn y la colitis ulcerosa. *R. gnavus* produce un antibiótico (RumA) que es activo frente a bacterias patógenas. La reducción en la población de *R. gnavus* en estos pacientes puede producir el crecimiento de bacterias perjudiciales (tales como bacterias patógenas) que a su vez está asociada a una respuesta inflamatoria. Por el contrario, determinadas bacterias asociadas a estas afecciones inflamatorias del intestino pueden producir antibióticos que inhiben bacterias beneficiosas, tales como especies de *Lactobacillus*. Al proporcionar el antibiótico (ya sea en forma purificada o como probiótico), los pacientes con estas afecciones pueden tratarse. Cualquier antibiótico producido por una bacteria descrito en este documento puede utilizarse para prevenir y/o tratar afecciones inflamatorias del intestino. El grupo RDP, el género representativo, o las especies de las bacterias enumeradas en las Tablas 3 y 5 se pueden utilizar para el diagnóstico, pronóstico y para el seguimiento de enfermedades de la presente invención. Por ejemplo, se observó un aumento en una *Moraxella osloensis* en la mucosa de Crohn en comparación con la mucosa de referencia. Esta información se obtuvo a partir de un clon secuenciado que se origina en la mucosa de un paciente con la enfermedad de Crohn. La secuencia de búsqueda de la base de datos RDP versión 8.1 (Cole J.R., Chai B., Marsh T.L., Farris R.J., Wang Q., Kulam S.A., Chandra S., McGarrell D.M., Schmidt T.M., Garrity G.M., Tiedje J.M. The Ribosomal Database Project (RDP-II): previewing a new autoaligner that allows regular updates and the new prokaryotic taxonomy. *Nucleic Acids Res.* 1 de enero de 2003, 31(1):442-3) (Véase, por ejemplo, [www.erdp8.cme.msu.edu/html](http://www.erdp8.cme.msu.edu/html)) indicaba que era un miembro del grupo *Pseudomonas* y afines a RDP, y análisis más preciso de la secuencia le asignó al género *Moraxella*. En el contexto de la presente invención, el grupo RDP solo, puede utilizarse como indicador de estado de la enfermedad. De este modo, en el ejemplo mencionado anteriormente, la clasificación de una bacteria como un miembro de *Pseudomonas* y afines al grupo RDP (independientemente de su clasificación en género o especie) es suficiente para indicar que el paciente que alberga la bacteria en su mucosa intestinal es más probable que padezca la enfermedad de Crohn, o que esté en regresión en una remisión temporal. Uno o más grupos pueden utilizarse para el diagnóstico. Por lo tanto, con respecto al ejemplo anterior, la determinación de una población microbiana de un paciente comprende tanto *Pseudomonas* como afines y las bacterias del grupo *Acidovorax* indican la existencia de la enfermedad de Crohn. Un análisis similar puede hacerse para todos los grupos RDP descritos en las tablas 3 y 5. Aunque no todas las permutaciones pueden darse a conocer en la aplicación, pueden seleccionarse de forma rutinaria en las Tablas 3, 5, y en las reivindicaciones adjuntas.

Para cualquier clon dado aislado de un paciente que se sospecha que padece la enfermedad de Crohn, colitis ulcerosa, o pouchitis, una identidad del 97% de la secuencia SSU del ARNr con una especie conocida es generalmente suficiente para que se clasifique como esta especie. Asimismo, aproximadamente el 95% de identidad es generalmente suficiente para la clasificación género y grupo RDP. La identidad se determinó utilizando el algoritmo BLAST (Tatusova, T. A., & Madden, T. L. (1999). BLAST 2 Sequences, a New Tool for Comparing Protein and Nucleotide Sequences. *FEMS Microbiology Letters*, 174, 247-250; Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids research*, 1 Sep. 1997, 25(17):3389-402; Wheeler, D. L., Chappey, C., Lash, A. E., Leipe, D. D., Madden, T. L., Schuler, G. D., et al. (2000). Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.*, 28(1), 10-14.).

La presente descripción proporciona procedimientos de diagnóstico, pronóstico, y/o seguimiento de la evolución de la enfermedad de Crohn, colitis ulcerosa o pouchitis, o en un paciente, que comprende: poner en contacto una muestra de tejido de la mucosa de colon que comprende ácido nucleico con una sonda de polinucleótidos que es específica para al menos una bacteria en condiciones eficaces para que dicha sonda se hibride específicamente con dicho ácido nucleico, y detectar la hibridación entre dicha sonda y dicho ácido nucleico, en los que un aumento, en comparación con una muestra normal de la mucosa, de una o más especies de bacterias o grupo RDP seleccionados de entre los siguientes indica la presencia de la enfermedad o el estado de la enfermedad: a) enfermedad de Crohn: *Moraxella sp.* del grupo *Pseudomonas*; *Comamonas sp.* del grupo *Acidovorax*; o *Cryseobacterium sp.* del grupo I de *Cytophaga* (en el que puede utilizarse el grupo RDP y/o el género y/o el género\_especie); b) colitis ulcerosa: *Moraxella sp.* de *Pseudomonas* y afines; *Comamonas sp.* del grupo *Acidovorax*; *Clostridium sp.* del grupo *Clostridium botulinum*; o *Enterococcus sp.* del grupo *Enterococcus* (en el que puede utilizarse el grupo RDP y/o género y/o género\_especie), o c) pouchitis (en comparación con la bolsa normal): *Ruminococcus sp.* del grupo *Clostridium\_Coccoides*; *Escherichia coli* y *Shigella sp.* de los grupos entéricos y afines; o *Fusobacterium sp.* del grupo *Fusobacteria* (en los que se puede utilizar el grupo RDP y/o el género y/o el género\_especie).

La presente descripción también proporciona procedimientos para el diagnóstico, pronóstico, y/o seguimiento de la evolución de las enfermedades de Crohn, colitis ulcerosa o pouchitis, que comprende: poner en contacto una muestra de tejido de la mucosa del colon que comprende ácido nucleico con una sonda de polinucleótidos que es específica para al menos una bacteria en condiciones eficaces para que dicha sonda se hibride específicamente con dicho ácido nucleico, y detectar la hibridación entre dicha sonda y dicho ácido nucleico, en los que una disminución, en comparación con una muestra de mucosa normal, de una o más bacterias seleccionadas de entre el grupo siguiente dichas bacterias indican la presencia de la enfermedad o el estado de la enfermedad: a) enfermedad de Crohn: *Bacteroides sp.* del grupo Bacteroides; *Propionibacterium sp.* del grupo Propionibacterium; o *Ruminococcus sp.* del grupo Clostridium\_Coccoides (en los que el grupo RDP y/o género y/o género\_especie puede utilizarse); b) la colitis ulcerosa: *Bacteroides sp.* del grupo Bacteroides; *Propionibacterium sp.* del grupo Propionibacterium; o *Ruminococcus sp.* del grupo Clostridium\_Coccoides (donde el grupo RDP y/o género y/o género\_especie puede utilizarse), c) pouchitis: *Bacteroides sp.* del grupo Bacteroides; *Propionibacterium sp.* del grupo de Propionibacterium (en los que se puede utilizar el grupo RDP y/o el género y/o el género\_especie).

La presente descripción también proporciona procedimientos para el diagnóstico, pronóstico y/o seguimiento de la evolución de las enfermedades de Crohn, colitis ulcerosa o pouchitis, en un paciente, que comprende: determinar la presencia de una o más de las siguientes bacterias en un tejido de la mucosa del colon en un paciente que padece la enfermedad de Crohn, colitis ulcerosa, o pouchitis: a) enfermedad de Crohn: *Morexella sp.* del grupo Pseudomonas; *Comamonas sp.* del grupo Acidovorax; o *Cryseobacterium sp.* del grupo I de Cytophaga (en el que puede utilizarse el grupo RDP y/o el género y/o el género\_especie); b) colitis ulcerosa: *Morexella sp.* de Pseudomonas y afines; *Comamonas sp.* del grupo Acidovorax; *Clostridium sp.* del grupo Clostridium botulinum; o *Enterococcus sp.* del grupo Enterococcus (en el que se puede utilizar el grupo RDP y/o el género y/o el género\_especie), c) pouchitis (en comparación con la bolsa normal): *Ruminococcus sp.* del grupo Clostridium\_Coccoides; *Escherichia coli* y *Shigella sp.* de los Entéricos y grupos afines; o *Fusobacterium sp.* del grupo Fusobacteria (en el que se puede utilizar el grupo RDP y/o el género y/o el género\_especie).

La presente descripción también proporciona procedimientos para el diagnóstico, pronóstico y/o el seguimiento de la evolución de las enfermedades de Crohn, colitis ulcerosa o pouchitis, que comprende: determinar la ausencia de una o más de las siguientes bacterias en un tejido de mucosa del colon de un paciente que padece la enfermedad de Crohn, colitis ulcerosa, o pouchitis: a) enfermedad de Crohn: *Bacteroides sp.* del grupo Bacteroides; *Propionibacterium sp.* del grupo Propionibacterium; o *Ruminococcus sp.* del grupo Clostridium\_Coccoides (en los que puede utilizarse el grupo RDP y/o el género y/o el género\_especie); b) la colitis ulcerosa: *Bacteroides sp.* del grupo Bacteroides; *Propionibacterium sp.* del grupo Propionibacterium; o *Ruminococcus sp.* del grupo Clostridium\_Coccoides (en los que se puede utilizar el grupo RDP y/o el género y/o el género\_especie), c) pouchitis: *Bacteroides sp.* del grupo Bacteroides; o *Propionibacterium sp.* del grupo Propionibacterium (en los que se puede utilizar el grupo RDP y/o el género y/o el género\_especie).

La presente descripción también proporciona un procedimiento para el diagnóstico, pronóstico y/o el seguimiento de la evolución de la enfermedad de Crohn o la colitis ulcerosa, en un paciente, que comprende: poner en contacto una muestra de lumen que comprende ácido nucleico con una sonda de polinucleótidos que es específica para al menos una bacteria en condiciones eficaces para que dicha sonda se hibride específicamente con dicho ácido nucleico, y detectar la hibridación entre dicha sonda y dicho ácido nucleico, en la que un aumento, en comparación con una muestra de lumen normal, de una o más bacterias seleccionadas de entre las siguiente indica la presencia de la enfermedad o el estado de la enfermedad: a) enfermedad de Crohn: *Bacteroides sp.* del grupo Bacteroides; o *Chryseobacterium sp.* del grupo I de Cytophaga (en los que puede utilizarse el grupo RDP y/o el género y/o el género\_especie), o b) colitis ulcerosa: *Bacteroides sp.* del grupo Bacteroides; o *Chryseobacterium sp.* del grupo I de Cytophaga (en los que se puede utilizar el grupo RDP y/o el género y/o el género\_especie).

La presente descripción también proporciona procedimientos para diagnosticar, pronosticar, y/o seguir la evolución de la enfermedad de Crohn o la colitis ulcerosa en un paciente, que comprende: poner en contacto una muestra del lumen que comprende ácido nucleico con una sonda de polinucleótidos que es específica para al menos una bacteria en condiciones eficaces para que dicha sonda se hibride específicamente con dicho ácido nucleico, y detectar la hibridación entre dicha sonda y dicho ácido nucleico, en donde una disminución, en comparación con una muestra normal de lumen, de *Acinetobacter sp.* o *Moraxella sp.* de la Pseudomonas y grupos afines indica que dicho paciente padece la enfermedad de Crohn o la colitis ulcerosa (en los que puede utilizarse el grupo RDP y/o el género y/o el género\_especie).

Sin más elaboración, se cree que un experto en la técnica puede, utilizando la descripción precedente, utilizar la presente invención en toda su extensión. Las siguientes formas de realización específicas preferidas deben interpretarse, por lo tanto, como meramente ilustrativas, y no restrictivas del resto de la descripción en modo alguno.

### Ejemplos

Recogida de muestras y extracción de ADN: Se recogieron muestras endoscópicas de tejido de la mucosa del ileon, terminal, ciego+colon ascendente, colon transversal, colon sigmoide y recto de los pacientes con IBD y pouchitis, así

como referencias sanas sometidas a la colonoscopia. Algunas de las muestras de tejido se lavaron en solución salina antes del análisis para eliminar las bacterias no adherentes (muestras lavadas frente a sin lavar). También se recogieron muestras de lumen conservadas con el endoscopio en el momento del procedimiento. Se tomaron huellas en las muestras para patrones bacterianos en la referencia 4, 2 UC, 4 CD y 3 pacientes con pouchitis, y 5 pacientes con bolsa sin pouchitis utilizando la metodología HLA. Las extracciones de ADN se realizaron utilizando el kit de suelo Bio101 de Qbiogene, Inc., Montreal, Quebec según las instrucciones del fabricante. Estos amplicones con HLA se agruparon, a continuación se clonaron y secuenciaron para identificar los componentes bacterianos que eran indicativos del estado de enfermedad.

10 Toma de huellas de la heterogeneidad de la longitud del amplicon (HLA): HLA es una técnica de toma de huellas bacteriana, véase Ritchie, N.J., *et al.*, Use of Length Heterogeneity PCR and Fatty Acid Methyl Ester Profiles to Characterize Microbial Communities in Soil. *Applied and Environmental Microbiology*, 2000. 66(4): p. 1668-1675; Suzuki, M., M.S. Rappe y S.J. Giovannoni, Kinetic bias in estimates of coastal picoplankton community structure obtained by measurements of small-subunit rRNA gene PCR amplicon length heterogeneity. *Applied and Environmental Microbiology* [Appl. Environ. Microbiol.]. HLA es un análisis basado en RCP que puede distinguir diferentes organismos basándose en variaciones naturales en la longitud de las secuencias de ADN ribosómico 16S. ADN purificado (10 ng) se amplió con RCP utilizando un cebador directo marcado con fluorescencia 27F (5'-[6FAM] AGAGTTTGATCCTGGCTCA G-3') y un cebador inverso sin marcar 338R' (5'-GCTGCCTCCCGTAGGAGT-3'). Ambos cebadores son muy específicos para las eubacterias. Alternativamente, los autores disponen de cebadores que amplían la región correspondiente en Archae véase Burggraf, S., T. Mayer, R. Amann, S. Schadhauer, C.R. Woese y K.O. Stetter, Identifying Members of the Domain Archaea with rRNA-Targeted Oligonucleotide Probes. *App. Environ. Microbiol.*, 1994. 60: págs. 3112-3119. Recientemente los autores han optimizado cebadores específicos para la ITS de hongos y demostraron que esa región del operón de ARNr tiene más información que la SSU ARNr véase Borneman, J. y J. Hartin, PCR primers that amplify fungal rRNA genes from Environmental Samples. *App. Environ. Microbiol.*, 2000. 66(10): págs. 4356-4360. Las reacciones se realizaron con mezclas de 50 µl (volumen final) que contenía 1 x tampón de RCP, albúmina de suero bovino al 0,6%, MgCl<sub>2</sub> 1,5 mM, cada trifosfato de desoxinucleótido a una concentración de 0,2 mM, cada cebador a una concentración de 0,2 µM, y 2 U de *Taq* ADN polimerasa. La desnaturalización inicial a 94°C durante 3 min fue seguida de 25 ciclos que consisten en la desnaturalización a 94°C durante 45 segundos, hibridación a 55°C durante 45 s, y ampliación a 72°C durante 2 min. Había una etapa de ampliación final que consta de 72°C durante 7 min. Las muestras con HLA se almacenaron a -20°C en la oscuridad hasta su utilización (generalmente menos de 1 semana).

Los productos HLA de RCP se separaron en el secuenciador capilar fluorescente SCE9610 (Spectrumedix LLC, State College, AP) y se analizaron con el paquete del programa informático GenoSpectrum. El programa informático convierte los datos de fluorescencia en electroferogramas. Los picos de los electroferogramas representan diferentes poblaciones de microflora de diferentes tamaños. Todos los datos de huellas se analizaron utilizando el programa informático (Interleave 1.0, BioSpherex LLC) que combina datos de varias pasadas, entrelaza los diferentes perfiles, normaliza los datos, y calcula índices de diversidad. Las áreas normalizadas de los picos se calcularon dividiendo cada una de las áreas de los picos por el área total del pico en ese perfil.

40 El análisis de huellas con HLA: La diversidad de las huellas con HLA se analizó utilizando índices de riqueza (R), equitatividad (E) y el índice de diversidad Shannon-Weaver (SW) en los grupos comparando IBD con las referencias, véase Hughes, J.B., *et al.*, Counting the Uncountable: Statistical Approaches to Estimating Microbial Diversity. *App. Environ. Microbiol.*, 2001. 67(10): págs. 4399-4406. Los parámetros relacionados con la IBD como la actividad de la enfermedad, y las partes histológicamente involucradas y no involucradas del íleon y del colon se compararon con los índices de diversidad.

Estas huellas se analizaron para determinar la agrupación global de las huellas con HLA en presencia o ausencia de IBD, tipos de IBD (CD, UC y pouchitis), actividad de la enfermedad, y las partes involucradas y no involucradas del íleon y colon (estado del tejido). Análisis Multidimensional de Reducciones [Análisis de Componentes Principales (PCA), análisis de coordenadas principales (PCO), Análisis de Correspondencia Canónica (CCA)] y el análisis de agrupamiento se realizó utilizando el paquete estadístico multivariable (MVSP), Kovach Computing Services, Gales, Reino Unido. Los siguientes análisis se llevaron a cabo. Generación de dendogramas por el método no ponderado de grupo par con medias aritméticas (UPGMA).

55 Análisis de Componentes Principales (PCA): PCA es uno de los métodos de ordenación más conocido y el más antiguo, descrito por primera vez por Karl Pearson (1901). Gráficamente, es una rotación de un enjambre de puntos de datos en el espacio multidimensional de manera que el eje más largo (el eje con la mayor varianza) es el primer eje de PCA, el segundo eje más largo perpendicular al primero es el segundo eje PCA y así sucesivamente. Los pocos primeros ejes PCA representan la mayor cantidad de variación en el conjunto de datos. Los primeros dos o tres ejes es de esperar que en general representen una gran proporción de la varianza, por ejemplo, 50-60% o más.

65 Análisis de Coordenadas Principales (PCO): PCO puede verse como una forma más general del PCA. PCO puede utilizar varias medidas diferentes de distancia o similitud. En general, las distancias o similitudes se miden entre los casos directamente, en lugar de las variables como en PCA. La principal ventaja de PCO es que se pueden utilizar muchos tipos diferentes de medidas de similitud o distancia. PCO está restringido a analizar distancias o similitudes

que son métricas y las distancias utilizadas deben poderse ver de alguna manera geométrica sensible por ejemplo, un triángulo.

Análisis de correspondencia canónica (CCA): En PCA y PCO, los datos se someten a algún tipo de manipulación matemática a fin de poner de manifiesto las tendencias más importantes. Estas tendencias se comparan entonces a menudo con otros datos relativos a las mismas muestras para determinar la relación entre los dos. Sin embargo, en el CCA, los datos están directamente relacionados. CCA es un método analítico multivariado de gradiente directo que ha llegado a utilizarse muy ampliamente en ecología.

Análisis por grupos: El análisis por grupos es un término utilizado para describir un conjunto de técnicas numéricas en las que el principal objetivo es dividir los objetos de estudio en grupos separados. Estos grupos se basan en las características de los objetos y se espera que los grupos tengan algún tipo de significación en relación con las cuestiones de investigación que se piden. El análisis por grupos se utiliza en muchas disciplinas científicas y se ha desarrollado una amplia variedad de técnicas para adaptarse a diferentes tipos de enfoques. Los más comúnmente utilizados son los métodos jerárquicos de aglomeración. Los métodos jerárquicos organizan los grupos en una jerarquía de modo que las relaciones entre los diferentes grupos son evidentes y los resultados se presentan en un diagrama tipo árbol, denominado dendrograma. Los métodos de aglomeración utilizados para crear un dendrograma empiezan combinando sucesivamente los objetos más similares hasta que todos están en un solo grupo jerárquico. Del mismo modo los dendrogramas se pueden crear utilizando el Método de grupo par no ponderado bien demostrado que utiliza medias aritméticas (UPGMA) y medios K.

Se identificaron supuestos patrones de huellas digitales con HLA (es decir, presencia o ausencia de determinados picos de ampliación) asociados a la presencia de IBD, a los tipos de enfermedades, a la actividad de la enfermedad y al estado de los tejidos. Con este objeto, los autores inspeccionarán a simple vista los histogramas de las huellas con HLA. Para determinar las correlaciones estadísticas de los picos con las variables relacionadas con IBD, los autores utilizaron también el análisis multivariante para grandes conjuntos de variables, es decir análisis discriminado y análisis de correspondencia canónica. Se emplearon también herramientas de minería de datos informatizados con algoritmos de reconocimiento de patrones supervisados y no supervisados. Estos incluyen C4.5, máquinas de vectores soporte y planos de autoorganización. Por lo tanto, estos análisis se utilizarán para determinar si las huellas con HLA pueden distinguir entre parámetros relacionados con IBD (presencia de enfermedad, tipo, actividad, estado del tejido) y determinar patrones específicos con HLA (presencia o ausencia de un pico o conjuntos de picos) asociados a IBD.

Secuenciación de clones con HLA: El producto de la RCP generado con cebadores utilizados para la huella con HLA se clonaron utilizando pGEM-T Easy Vector System II (Promega Corp., Madison, WI). Los clones se seleccionaron evaluando inserciones que utilizan la complementación alfa con X-Gal (5-bromo-4-cloro-3-indoil-B-D-galactopiranosido) e IPTG (isopropil-BD-tiogalactopiranosido). Las inserciones se secuenciaron utilizando la química del colorante terminador *Taq* y los productos de secuenciación se separaron en un secuenciador capilar fluorescente SCE9610.

Análisis de los datos de clones con HLA: Las secuencias anteriores de clones con HLA se compararon con las secuencias en la base de datos RDP para evaluar patrones de microflora utilizando un nuevo programa (CloneID 1.0, BioSpherex LLC). El algoritmo utiliza básicamente Megablast para comparar los datos de la secuencia de clones con la base de datos RDP y compila una tabla con los números de RDP para correlacionar la identificación con un esquema de clasificación jerárquica. De estos mismos clones con HLA se tomaron sus huellas digitales para determinar el tamaño HLA empírico y se correlacionaron con la huella con HLA original de la muestra utilizando un segundo programa (CloneMatch 1.0, BioSpherex, LLC).

#### Resultados de análisis de Crohn (CD) y colitis ulcerosa (UC)

Aunque las huellas con HLA varían cualitativa y cuantitativamente entre los individuos, existen patrones de diagnóstico muy diferentes que se pueden ver procedentes del análisis de muestras de tejidos combinados (mucosa) y de lumen. La figura 1 es un histograma compilado de la media de la huella con HLA de todas las muestras de Crohn y muestras de referencia (es decir, todos los individuos y todas las posiciones) que muestra las longitudes de amplificación en pares de bases (pb) en el eje x y las abundancias relativas en el eje y. Las muestras tisulares combinadas de las referencias (barras blancas) tenían el perfil con HLA muy distinto que difieren drásticamente de las muestras de lumen de referencia (barras negras) lo que indican que hay una población microflora distinta que se adhiere a la mucosa como un biofilme. En cambio no había una clara diferenciación entre la microflora del lumen y del tejido en la enfermedad de Crohn lo que indica una disbiosis drástica en la que muchas de las especies de bacterias normalmente en el lumen se encuentran en el biofilme. Por lo tanto, existe mucha más superposición entre los amplicones con HLA de tejido de Crohn (barras gris claro) y el lumen de Crohn (barras gris oscuro) con lumen de referencia (barras negras). Existen amplicones con HLA para diagnóstico que se producen principalmente en las muestras de lumen de referencia y el tejido de Crohn (es decir, en 333,0 pb, 334,3 pb y 338,6 pb). Además, existen algunos amplicones con HLA que son exclusivos para el tejido de Crohn (es decir 310,9 pb y 313,4 pb), pero por término medio, representan una proporción pequeña de la población de microflora.

Del mismo modo, parece que existen disbiosis en la colitis ulcerosa (UC) ya que los perfiles con HLA de tejido con UC y lumen de UC son similares para el lumen de referencia al perfil con HLA de las referencias que son muy distintas (Figura 2). Por lo tanto, había mucha más superposición entre los amplicones con HLA de tejido con UC (barras gris claro) y lumen con UC (barras gris oscuro) con lumen de referencia (barras negras). Algunos de los amplicones con HLA de diagnóstico que se observaron en CD (véase más arriba) son los mismos amplicones que son para diagnóstico en UC (es decir, a 333,0 pb, 334,3 pb y 338,6 pb). Además, hay diferentes amplicones de HLA que se producen sólo en el tejido de UC (es decir, 334,6 pb).

Cuando las diferencias de carácter medio para los perfiles con HLA de referencias, CD y UC se examinaron utilizando el Análisis de Coordenadas Principales (PCO), pueden observarse patrones drásticos de agrupamiento para UC y CD que son distintos de las muestras de referencia (figura 3). Los autores observan claramente distinta agrupación de perfiles con HLA de referencia en el 1<sup>er</sup> cuadrante, grupos con UC en el 3<sup>er</sup> 4<sup>o</sup> cuadrantes, y CD se agrupan principalmente en el 2<sup>o</sup> cuadrante. También es importante observar que las muestras de lumen se agrupan en el 3<sup>er</sup> y 4<sup>o</sup> cuadrantes asociado a la UC. Hay también varios perfiles con HLA de Crohn que se agrupan en estos 3<sup>er</sup> y 4<sup>o</sup> cuadrantes lo que sugiere que hay variación en la microflora del tejido de Crohn y que, en muestras específicas, estos perfiles con HLA son similares a los de la UC.

El PCA y los análisis de correspondencia canónica demuestran una agrupación similar de referencias sanas independientes de los pacientes con EC y UC. Los dendogramas producidos con la agrupación UPGMA utilizando una medida de distancia Jacard también muestran los mismos patrones generales que el análisis PCO.

Se han clonado y secuenciado amplicones con HLA mezclados de muestras de UC, CD y de referencias sanas y estos datos de secuencia se utilizaron para identificar las especies bacterianas asociadas a cada enfermedad. La tabla 1 resume los principales grupos bacterianos basado en el esquema de clasificación de RDP que se producen con una frecuencia de más del 5% de la población de microflora. Los datos apoyan los perfiles con HLA en la microflora que se encuentra en la superficie de la mucosa de tejido con CD y UC se asemejan a la composición de la microflora de lumen en individuos sanos y que esta composición difiere de la composición de la microflora de la mucosa de las referencias. En concreto, los miembros de las Pseudomonas, tales como *Moraxella sp.* y miembros del grupo Acidovorax tal como *Comoamonas sp.* están asociados con el lumen de referencia, lumen con CD, mucosa con CD, lumen con UC y la mucosa con UC. Además, los miembros del grupo Cytophaga tal como *Chryseobacterium balustinum* están asociados a la mucosa con CD, lumen con CD y lumen con UC. Por último, los miembros tanto del grupo Clostridium (*Clostridium paraputrificum*) como Enterococcus (*Enterococcus hirae*) también están asociados a la mucosa con UC. Se observa asimismo una disminución cuantitativa en el grupo de Bacteroides en la mucosa con UC y la mucosa con CD en comparación con la mucosa de referencia. En resumen, podemos concluir que existen especies bacterianas que están asociadas en el biofilme con CD y el biofilme con UC que se encuentran normalmente en el lumen y que esto indica disbiosis grave.

#### Resultados de análisis de pouchitis

La figura 4 es un histograma compilado a partir de la media de la huella de HLA de todas las muestras con pouchitis (AP) y las muestras de bolsa normal (NP), es decir muestras de pacientes con pouchitis activa (AP) y pacientes con una bolsa pero que son normales en el examen (NP). Como se aprecia en CD y UC, las muestras mezcladas de mucosa con NP (barras blancas) tenían el perfil con HLA muy distinto que difería drásticamente de las muestras de mucosa con NP (barras negras) que indican que existe una población de microflora distinta que se adhiere a la mucosa como un biofilme. Además, los perfiles de amplicón con HLA de las muestras de NP fueron diferentes de los pacientes de referencia sanos que no tenían una bolsa. Existen amplicones con HLA para diagnóstico que se producen predominantemente en las muestras de mucosa con AP (es decir, a 309,2 pb, 310,0 pb, 310,9 pb, 331,2 pb, 350,2 pb, y 359,6 pb) que no son los amplicones con HLA predominantes de diagnóstico en CD y UC. Por lo tanto, la disbiosis en la pouchitis parece ser muy diferente que en CD y UC y puede implicar diferentes patologías. Además, los componentes reales de la población en la enfermedad varían de un individuo a otro. Por ejemplo, los amplicones con HLA en 309,2 pb, 310,0 pb, 310,9 pb son los principales componentes en un paciente, pero sólo son componentes menores de otros. Es importante destacar que los pacientes con bolsa normal tienen contenido anormal de microflora en el biofilme de la mucosa y estos pacientes pueden estar continuamente en un estado semipatológico.

Cuando las diferencias medias de carácter para los perfiles con HLA de muestras de NP y AP se examinaron utilizando el análisis de coordenadas principales (PCO), un patrón de agrupamiento general puede observarse para NP en el centro de la gráfica que es distinto de los tres grupos de muestras con AP (figura 5). Curiosamente, cada uno de estos grupos de AP son de pacientes independientes lo que confirma que los pacientes con pouchitis presentan mucha más variación en la microflora en el biofilme de la mucosa. Cabe señalar que el grupo separado encontrado en el eje Y por encima del grupo de bolsa normal es el paciente que presentaba los amplicones con HLA distintos a 309,2 pb, 310,0 pb y 310,9 pb. El grado de actividad de la enfermedad se puede reflejar en la medida de disbiosis representada en el diagrama de PCO. Además, el patrón es coherente si las muestras se han lavado o no se han lavaron en solución salina como es el caso en las muestras de CD y de UC.

Se han clonado y secuenciado amplicones con HLA mezclados de muestras de mucosas con NP y AP y estos datos de secuencia se utilizaron para identificar las especies bacterianas asociadas a cada enfermedad. La tabla 1 resume los principales grupos bacterianos basándose en el esquema de clasificación de RDP que se produce a una frecuencia de más del 5% de la población de microflora. Los datos apoyan los perfiles con HLA porque las microfloras que se encuentran en la superficie de la mucosa tanto de tejido con AP como con NP son diferentes de las que se encuentran en individuos sanos y éstos no reflejan la microflora encontrada en lumen normal como la encontrada en CD y UC. Específicamente, los miembros del grupo Clostridium (es decir *Clostridium paraputrificum*), los miembros de entéricos (es decir, *E. coli* y *Shigella sp.*) y los miembros del grupo Streptococcus (es decir, *Streptococcus brevis*) se encuentran asociados a la mucosa de pacientes con NP. Por otra parte, la microflora asociada a la mucosa en pacientes con AP era muy diversa y se diferenciaba de la de los pacientes con NP. Específicamente, los autores observaron que los miembros de los Entéricos (es decir *E. coli* y *Shigella sp.*) y del grupo Fusobacterium (es decir, *Fusobacterium varium*) estaban asociados a la mucosa en los pacientes con AP y que había una pérdida drástica de miembros del grupo Streptococci (es decir *Streptococcus brevis*). Además, una especie diferente *Ruminococcus (Ruminococcus obeum)* estaba asociada a pacientes con AP, pero no está claro que esta cepa diferente contribuyese a la patología. En resumen, parece que tanto los pacientes con NP como con AP tienen disbiosis en comparación con las referencias normales, y que hay una pérdida dramática de estreptococos en los pacientes con AP. Por otra parte, parece ser específico del paciente (véase figura 5) huellas con HLA, lo que sugiere una variación significativa en la microflora entre los pacientes.

20 Correlación de amplicones con HLA y microflora

Se ha correlacionado el tamaño del amplicón con HLA determinado experimentalmente de clones con las identificaciones obtenidas a partir de esta secuenciación. Por ejemplo, se han marcado los amplicones principales en las huellas con HLA de pouchitis y bolsa normal del histograma en la figura 6. A continuación, se utilizan esta información para correlacionar qué especies bacterianas están en los picos de diagnóstico de los perfiles con HLA.

Los ejemplos anteriores pueden repetirse con éxito similar sustituyendo los reactivos descritos genérica o específicamente y/o las condiciones de operación de la presente invención por los utilizados en los ejemplos anteriores.

30 Tabla 1 Aumento o disminución de amplicones con huellas con HLA superior al 5% en mucosa de Crohn frente a mucosa de referencia

Tamaño de amplicón (pb)	333,0	334,3	336,2	337,6	338,6	340,2	341,9	342,6	343,6	347,3	347,9	349,3	351,6	358,3
Mucosa de referencia			2%	1%		6%	6%	12%	1%	4%	11%	21%	16%	5%
Mucosa de referencia	20%	7%	2%	3%	2%		3%	5%	6%		1%	31%	1%	
Aumento en la mucosa de Crohn	333,0	334,3			338,6				343,6					
Disminución en la mucosa de Crohn						340,2	341,9	342,6		347,9	347,9	349,3	351,6	358,3

35 Tabla 2 Aumento o disminución de amplicones con huellas con HLA mayores del 5% en la mucosa con colitis ulcerosa frente a la mucosa de referencia

Tamaño de amplicón (pb)	333,0	334,3	340,2	341,9	342,6	347,9	349,3	351,6	358,3
Mucosa de referencia			6%	6%	12%	11%	21%	16%	5%
Mucosa con UC		39%	13%	9%	3%	7%	7%		
Aumento en la mucosa con UC	333,0	334,3	340,2						
Disminución en la mucosa con UC				341,9	342,6	347,9	349,3	351,6	358,3

TABLA 3 Especies bacterianas asociadas a pouchitis

GRUPO RDP	Ejemplo de género especie	Mucosa de referencia	Bolsa normal	Mucosa de pouchitis
GRUPO BACTEROIDES (2.15.1.2.8)	Bacteroides vulgatus	0,38		0,28
GRUPO BACTEROIDES (2.15.1.2.7)	Bacteroides distasonis	0,08		
GRUPO PROPIONIBACTERIUM (2.30.1.10.1)	Propionibacterium acnes	0,06		
GRUPO CLOSTRIDIUM COCCOIDES (2.30.4.1.4)	Ruminococcus gnavus	0,09	0,13	0,08
GRUPO CLOSTRIDIUM COCCOIDES (2.30.4.1.1)	Ruminococcus obeum			0,14
PSEUDOMONAS Y AFINES (2.28.3.13.1.6)	Acinetobacter junii			
PSEUDOMONAS Y AFINES (2.28.3.13.1.1)	Moraxella osloensis	0,10		
GRUPO ACIDOVORAX (2.28.2.9.4.14)	Comamonas sp	0,41		
GRUPO ACIDOVORAX (2.28.2.9.4.1)	Comamonas terrigena	0,14		
GRUPO CYTOPHAGA I (2.15.1.3.6)	Chryseobacterium balustinum	0,15		
GRUPO CLOSTRIDIUM BOTULINUM (2.30.9.2.11.14)	Clostridium paraputrificum		0,08	
GRUPO ENTEROCOCCUS (2.30.7.20)	Enterococcus hirae			
ENTÉRICOS Y AFINES (2.28.3.27.2)	E. coli/Shigella sp		0,18	0,11
GRUPO STREPTOCOCCI (2.30.7.21.6)	Streptococcus bovis		0,34	
GRUPO FUSOBACTERIA (2.29.5)	Fusobacterium varium			0,12

Tabla 4 Aumento o disminución de amplicones con huellas con HLA mayores del 5% en mucosa con pouchitis frente a bolsa normal

Tamaño de amplicón (pb)	309,2	310,0	310,9	329,8	331,2	340,2	341,9	342,6	349,3	350,2	356,6	357,5	359,6
Bolsa normal	1%			5%		2%	11%	11%	17%		32%	5%	1%
Pouchitis	17%	5%	5%		7%	7%	5%	2%	9%	6%	19%	1%	6%
Aumento en pouchitis	309,2	310,0	310,9		331,2	340,2				350,2	356,6		359,6
Disminución en pouchitis				329,8			341,9	342,6	349,3			357,5	



TABLA 5 Especies bacterianas asociadas a la enfermedad de Crohn y la colitis ulcerosa

GRUPO RDP	Ejemplo de género especie	Mucosa de referencia	Lumen de referencia	Mucosa de Crohn	Lumen de Crohn	Mucosa de UC	Lumen de UC
GRUPO BACTEROIDES (2.15.1.2.8)	<i>Bacteroides vulgatus</i>	0,38			0,19		0,07
GRUPO BACTEROIDES (2.15.1.2.7)	<i>Bacteroides distasonis</i>	0,08					
GRUPO PROPIONIBACTERIUM (2.30.1.10.1)	<i>Propionibacterium acnes</i>	0,06					
GRUPO CLOSTRIDIUM COCCOIDES (2.30.4.1.4)	<i>Ruminococcus gnavus</i>	0,09					
GRUPO CLOSTRIDIUM COCCOIDES (2.30.4.1.1)	<i>Ruminococcus obeum</i>						
PSEUDOMONAS Y AFINES (2.28.3.13.1.6)	<i>Acinetobacter junii</i>		0,10				
PSEUDOMONAS Y AFINES (2.28.3.13.1.1)	<i>Moraxella osloensis</i>		0,41	0,29	0,15	0,10	0,30
GRUPO ACIDOVORAX (2.28.2.9.4.14)	<i>Comamonas</i> sp		0,14	0,13	0,12	0,12	0,14
GRUPO ACIDOVORAX (2.28.2.9.4.1)	<i>Comamonas terrigena</i>		0,15	0,29	0,23	0,39	0,19
GRUPO CYTOPHAGA I (2.15.1.3.6)	<i>Chryseobacterium balustinum</i>			0,09	0,08		0,11
GRUPO CLOSTRIDIUM BOTULINUM (2.30.9.2.1.1.14)	<i>Clostridium paraputrificum</i>					0,06	
GRUPO ENTEROCOCCUS (2.30.7.20)	<i>Enterococcus hirae</i>					0,06	
ENTERICOS Y AFINES (2.28.3.27.2)	<i>E. coli/Shigella</i> sp						
GRUPO STREPTOCOCCI (2.30.7.21.6)	<i>Streptococcus bovis</i>						
GRUPO FUSOBACTERIA (2.29.5)	<i>Fusobacterium varium</i>						

TABLA 6

REFERENCIA DE RDP	SEC. ID. nº	GÉNERO ESPECIE
2.28-3.13.1.6	14	Acinetobacter Junii
2.15.1.2.7	1	Bacteroides distasonis
2.15.1.2.8	2	Bacteroides fragilis
2.15.1.2.8	3	Bacteroides ovatus
2.15.1.2.8	4	Bacteroides vulgatus
2.15.1.3.6	17	Bergeyella zoohelcum
2.15.1.3.6	18	Chryseobacterium balustinum str. SBR1044
2.15.1.3.6	19	Chryseobacterium balustinum str. SBR2024
2.28.2.9.4.1	16	Comamonas terrigena
2.28.2.9.4.14	15	Comamonas sp.
2.28.3.13.1.1	11	Moraxella cuniculi
2.28.3.13.1.1	12	Moraxella lacunata
2.28.3.13.1.1	13	Moraxella osloensis
2.28.3.27.2	25	Escherichia coli
2.28.3.27.2	26	Salmonella bovis
2.28.3.27.2	27	Shigella boydii
2.28.3.27.2	28	Shigella dysenteriae
2.28.3.27.2	29	Shigella flexneri
2.29.5	34	Fusobacterium alocis
2.29.5	35	Fusobacterium nucleatum
2.29.5	36	Fusobacterium varium
2.30.1.10.1	5	Propionibacterium acnes
2.30.4.1.1	10	Clostridium sp.
2.30.4.1.4	6	Clostridium nexile
2.30.4.1.4	7	Eubacterium formicigenerans
2.30.4.1.4	8	Ruminococcus gnavus
2.30.4.1.4	9	Pares Ruminacoccus
2.30.7.20	21	Enterococcus cecorum
2.30.7.20	22	Enterococcus columbae
2.30.7.20	23	Enterococcus hirae
2.30.7.20	24	Tetragenococcus halophilus
2.30.7.2.1.6	30	Streptococcus bovis
2.30.7.21.6	31	Streptococcus infantarius
2.30.7.21.6	32	Streptococcus salivarius
2.30.7.21.6	33	Streptococcus thermophilus
2.30.9.2.11.4	20	Clostridium paraputrificum

5

**Listado de secuencias**

<110> GEORGE MASON UNIVERSITY

10 <120> COMPOSICIONES Y PROCEDIMIENTOS PARA EL DIAGNÓSTICO DE TRASTORNOS DEL COLON

<130> GMU-001-WO

<140>

15 <141>

<150> 60/646,592

<151> 2005-01-26

20 <150> 60/673,771

<151> 2004-11-01

<160> 38

25 <170> PatentIn versión 3.3

<210> 1

<211> 1489  
 <212> ARN  
 <213> Bacteroides distasonis

5 <220>  
 <221> modified\_base  
 <222> (39)..(39)  
 <223> n es a, c, g, o u

10 <220>  
 <221> modified\_base  
 <222> (99) .. (99)  
 <223> n es a, c, g, o u

15 <220>  
 <221> modified\_base  
 <222> (101)..(101)  
 <223> n es a, c, g, o u

20 <220>  
 <221> modified\_base  
 <222> (108)..(108)  
 <223> n es a, c, g, o u

25 <220>  
 <221> modified\_base  
 <222> (235)..(235)  
 <223> n es a, c, g, o u

30 <220>  
 <221> modified\_base  
 <222> (482)..(482)  
 <223> n es a, c, g, o u

35 <220>  
 <221> modified\_base  
 <222> (656)..(656)  
 <223> n es a, c, g, o u

40 <220>  
 <221> modified\_base  
 <222> (932)..(932)  
 <223> n es a, c, g, o u

45 <220>  
 <221> modified\_base  
 <222> (936)..(936)  
 <223> n es a, c, g, o u

50 <220>  
 <221> modified\_base  
 <222> (1160)..(1160)  
 <223> n es a, c, g, o u

55 <400> 1

ES 2 391 744 T3

caauuuuaac aacgaagagu uugauccugg CUCaggauna acgcuagcga caggcuuaac 60  
acaugcaagu cgaggggcac gcgcgrguag Caaauaccng ngcuggcnac cggcgcacgg 120  
gugaguaacg cguaugcaac uugccuauca Gaggggggaau acccggcgaa agucggacua 180  
auaccgcaug aagcagggau cccgcauggg Aauuuugcu aaagauucau cgcunauaga 240  
uaggcaugcg uuccauuagg caguuggcgg Gguaacggcc caccaaaccg acgauaggaua 300  
ggguuucuga gaggaagguc ccccacauug Guacugagac acggaccaa cuccuacggg 360  
aggcagcagu gaggaauuuu ggucaauggc Cgagagcgug aaccagccaa gucgcgugag 420  
ggauagaaggu ucuauggauc guaaaccucu Uuuuaaaggg aauaaagugc gggacguguc 480  
cnguuuugua uguaccuuau gaauaaggau Cggcuaaacuc cgugccagca gccgcgguaa 540  
uacggaggau ccgagcguaa uccggauuuu Uuggguuuua agggugcgua ggcggccuuu 600  
uaagucagcg gugaaagucu guggcucaac Cauagaauug ccguugaaac uggggngcuu 660  
gaguauguuu gaggcagcgg gaaugcgugg Uguagcgugg aaaugcauag auaucacgca 720  
gaaccccgau ugcgaaaggca gccugccaag Ccauuacuga cgcugaugca cgaaagcgug 780  
gggaucaaac aggauuagau acccugguag Uccacgcagu aaacgaugau cacuagcugu 840  
uugcgauaca cuguaagcgg cacagcgaaa Gcguaaagug auccaccugg ggaguaagcc 900  
ggcaaccggg aaacucuaag gaauugacgg Gngccngcac aagcggaggga acaugugguu 960  
uaauucgaug auacgcgagg aaccuuacc Ggguuugaac gcauucggac cgagguggaa 1020  
acaccuuuuc uagcaauagc cguuugcgag Ggugcugcaug guugucguca gcucgugccg 1080  
ugaggugucg gcuaaagugc cauaacgagc Gcaaccuug ccacuaguua cuaacagguu 1140  
aggcugagga cucugggugn acugccagcg uaagcugcga ggaaggcggg gaugacguca 1200  
aaucagcacg gcccuuacau ccggggcgac acacguguua caauggcgug gacaaagga 1260  
ggccaccugc cgacagggag cgaauccca aaccacgucu caguucggau cggagucugc 1320  
aaccgcacuc cgugaagcug gauucgcuag uaaucgcgca ucagccaugg cgcggugaau 1380  
acguuuccgg gccuuuaca caccgccgu caagccaugg gagccggggg uaccugaagu 1440  
ccguaaccga aaggauccgg cuaggguaaa acuggugacu ggggcuaag 1489

5 <210> 2  
<211> 1533  
<212> ARN  
<213> Bacteroides fragilis

10 <400> 2

ES 2 391 744 T3

uuacaacgaa gaguuugauc cuggcucagg augaacgcua gcuacaggcu uaacacaugc 60  
aagucgaggg gcaucaggaa gaaagcuugc uuucuuugcu ggcgaccggc gcaccggguga 120  
gnaacacgna uccaaccugc ccuuuacucg gggauagccu uucgaaagaa agauuaauac 180  
ccgauagcau aaugauuccg caugguuua uauuuuaagg auuccgguaa aggauagggga 240  
ugcguuccau uagguuguug g'ugagguaac ggcucaccaa gccuucgaur gauagggguu 300  
cugagaggaa ggucccccac auuggaacug agacacgguc caaacuccua cgggagggcag 360  
cagugaggaa uauuggucaa ugggcgcuag ccugaaccag ccaaguagcg ugaaggaua 420  
aggcucuaug ggucguaaac uucuuuuua uaagaauaaa g'ugcaguaur uauacuguu 480  
uguauguuu auaugaauaa ggauccgcu acuccgugcc agcagcccg'c guaaucggga 540  
ggauccgagc guuauccgga uuuuuugggu uuaaaggag'c cguaggugga cugguaaguc 600  
aguugugaaa guuugcg'cu caaccguaaa auugcagcug auacugucag ucuugaguac 660  
aguagaggug ggcggaauc gugguguagc ggugaaaugc uuagauauca cgaagaacuc 720  
cgauugcgaa ggcagcucac ug'gacugcaa cugacacuga ugcucgaaag uguggguauc 780  
aaacaggauu agauacc'ug guaguccaca caguaaacga ugaauacuc'g cuguuugcga 840  
uauacaguaa g'cgcccaagc gaaagcauuu aguuuuccac cuggggg'gua c'ccggcaac 900  
ggugaaacuc aaaggaaucg acgggggccc gcacaagcgg aggaacaugu gguuuuuuuc 960  
gaugauacgc gaggaaccuu acccg'gcu aaauugcagu ggaauaugu g'gaaacaugu 1020  
cagugagcaa uacccg'cu gaaggugcug caugguuguc g'ucagcuc'cu gccgugaggu 1080  
gucggcuuaa g'ugccauaac gagcgcaacc cuuauuuua guuacuaaca g'guuauugcug 1140  
aggacucua'g agagacugcc g'ucgnaagau g'ugaggaagg uggggaugac gucaaaucag 1200  
cacggccuu acguccgggg cuacacacgu guuacaau'g gggguacaga aggcagcuag 1260  
cgggugaccg uaugcuauac ccaaaauc'cu cucucaguuc ggau'cgaagu cugcaacc'cg 1320  
acuucgugaa gcuggauucg cuaguauucg cgcaucagcc acggcgcgg'g gaauacguuc 1380  
ccggg'ccuug uacacacc'cg ccgucaagcc au'gggagccg gggguacc'ug aaguacguaa 1440  
ccgcaaggau c'guc'cuaggg uaaaacug'gu gacugggg'cu aagucg'uaac aagguagccg 1500  
uaccggaagg ugcggcugga acaccuccuu ucu 1533

5 <210> 3  
<211> 1468  
<212> ARN  
<213> Bacteroides ovatus

10 <220>  
<221> modified\_base  
<222> (204)..(207)  
<223> n es a, c, g, o u

15 <220>  
<221> modified\_base  
<222> (339)..(339)  
<223> n es a, c, g, o u

20 <220>  
<221> modified\_base  
<222> (439) .. (439)  
<223> n es a, c, g, o u

# ES 2 391 744 T3

<220>  
<221> modified\_base  
<222> (530) .. (530)  
<223> n es a, c, g, o u  
5  
<220>  
<221> modified\_base  
<222> (544) .. (544)  
<223> n es a, c, g, o u  
10  
<220>  
<221> modified\_base  
<222> (700) .. (700)  
<223> n es a, c, g, o u  
15  
<220>  
<221> modified\_base  
<222> (739)..(739)  
<223> n es a, c, g, o u  
20  
<220>  
<221> modified\_base  
<222> (788)..(788)  
<223> n es a, c, g, o u  
25  
<220>  
<221> modified\_base  
<222> (953)..(953)  
<223> n es a, c, g, o u  
30  
<220>  
<221> modified\_base  
<222> (962)..(962)  
<223> n es a, c, g, o u  
35  
<220>  
<221> modified\_base  
<222> (988)..(988)  
<223> n es a, c, g, o u  
40  
<220>  
<221> modified\_base  
<222> (1031)..(1031)  
<223> n es a, c, g, o u  
45  
<220>  
<221> modified\_base  
<222> (1084)..(1084)  
<223> n es a, c, g, o u  
50  
<400> 3

ES 2 391 744 T3

augaagaguu ugauccuggc ucaggaugaa cgcuaagcuac aggcuaaaca caugcaaguc 60  
 gaggggagc auuuuaguuu gcuvgcaaac ugaagauggc gaccggcgca cgggugagua 120  
 acacguaucc aaccugccga uaacuccggg auagccuuuc gaaagaaaga uuaauaccgg 180  
 aurgyauayg aacaucgcgau gaunnnnuua uuaaagaauu ucgguuuucg auggggaguc 240  
 guuccauuag uuuguuuggc ggguaacggc ccaccaagac uacgauggau agggguucug 300  
 agagggaaggu cccccacauu ggaacugaga cacgguccna acuccuacgg gagggcagcag 360  
 ugaggaaauu uggucaaugg gcgagagccu gaaccagcca aguagcguga aggaugaagg 420  
 cucuaugggu cguaaacunc uuuuauaugg gaauaaagug uuccacgugu ggaauuuugu 480  
 auguaccuaa ugaauaagga ucggcuaacu ccgugccagc agccgggun auaccggagga 540  
 uccnagcguu auccggauuu auuggguuuu aagggagcgu aggggguug uuaagucagu 600  
 ugugaaaguu ugcggcucaa ccguaaaauu gcaguugaaa cuggcagucu ugaauacagu 660  
 agaggugggc ggaauucgug guguagcggg gaaaugcuun gauaucacga agaacuccga 720  
 uugcgaaggc agcucacung acuguuacug acacugaugc ucgaaagugu ggguaucaaa 780  
 caggauunga uaccuggua guccacacag uaaacgauga auacucgug uuugcgauau 840  
 acaguaagcg gccaaagcga agcauuuagu auuccaccug gggaguacgc cggcaaccgu 900  
 gaaacucaaa ggaauugacg ggggcccgca caagcggagg aacauguggu uuaauucgau 960  
 gnuacgcgag gaaccuuacc cgggcuuuaa uugcawcwg auaauwggg aacwruauag 1020  
 ccgyaaggca nuugugaagg ugcugcaugg uugucgucag cucgugccgu gaggugucgg 1080  
 cuunagugcc auaacgagcg caaccuuau cuuaguuac uaacaggua ugcugaggac 1140  
 ucuagagaga cugccgucgu aagaugugag gaaggugggg augacgucaa auacgcacgg 1200  
 cccuacguc cggggcuaca cacguguuac aauggggggu acagaaggcr gcuaaccuggy 1260  
 gacaggauugc uaauccaaa aaccucucuc aguucggauc gaagucugca acccgacuuc 1320  
 gugaagcugg auucgcuagu aaucgcgcau cagccauggc gcggugaaua cguuaccggg 1380  
 ccuuguaacac accgcccguc aagccaugaa agccgggggu accugaagua cguaacccga 1440  
 aggagcgucc uaggguaaaa cugguaau 1468

5 <210> 4  
 <211> 1524  
 <212> ARN  
 <213> Bacteroides vulgatus

10 <220>  
 <221> modified\_base  
 <222> (35) .. (35)  
 <223> n es a, c, g, o u

15 <220>  
 <221> modified\_base  
 <222> (97) .. (97)  
 <223> n es a, c, g, o u

20 <220>  
 <221> modified\_base  
 <222> (99)..(99)  
 <223> n es a, c, g, o u

25 <220>  
 <221> modified\_base

ES 2 391 744 T3

- <222> (395)..(395)
- <223> n es a, c, g, o u
  
- <220>
- 5 <221> modified\_base
- <222> (481) .. (481)
- <223> n es a, c, g, o u
  
- <220>
- 10 <221> modified\_base
- <222> (544)..(544)
- <223> n es a, c, g, o u
  
- <220>
- 15 <221> modified\_base
- <222> (744)..(744)
- <223> n es a, c, g, o u
  
- <220>
- 20 <221> modified\_base
- <222> (932)..(932)
- <223> n es a, c, g, o u
  
- <220>
- 25 <221> modified\_base
- <222> (1204)..(1204)
- <223> n es a, c, g, o u
  
- <220>
- 30 <221> modified\_base
- <222> (1255)..(1255)
- <223> n es a, c, g, o u
  
- <220>
- 35 <221> modified\_base
- <222> (1499)..(1499)
- <223> n es a, c, g, o u

<400> 4

40

uauuacaau	aagaguuga	uccuggcuca	ggaunaacgc	uagcuacagg	cuuaacacau	60
gcaagucgag	gggcagcaug	gucuuagcuu	gcuaaagnca	uggcgaccgg	cg cacgggug	120
aguaaacag	auccaaccug	ccgucuauc	uuggacagcc	uucugaaagg	aa gauuaaua	180
caagauggca	ucaugagucc	gcauguucac	augauuaaag	guauuccggu	ag acgauggg	240
gaugcguucc	auuagauagu	aggcggggua	acggcccacc	uagucuucga	ug gauagggg	300
uucugagagg	aagguccccc	acauuggaac	ugagacacgg	uccaaacucc	ua cgggaggc	360
agcagugagg	aaauugguc	aaugggcgag	agccngaacc	agccaaguag	cg ugaaggau	420
gacugcccu	uggguugua	acuucuuua	uaaaggaau	aagucgggua	ug gauacccg	480
nuugcaugua	cuuuaugaau	aaggaucggc	uaacuccgug	ccagcagccg	cg guaauacg	540
gagnauccga	gcguuaucgg	gauuuauugg	guuuaaaggg	agcguagaug	ga uguuuuag	600
ucaguuguga	aaguuugcgg	cucaaccgua	aaauugcagu	ugauacugga	ua ucuugagu	660
gcaguugagg	caggcggaau	ucguggugua	gcguguaaa	gcuuagauau	ca cgaagaac	720
uccgauugcg	aaggcagccu	gcunagcugc	aacugacauu	gaggcucgaa	ag rugugggua	780
ucaaacagga	uuagauaccc	ugguagucca	cacgguaaac	gaugaauacu	cg cuguuugc	840



ES 2 391 744 T3

gauauacugc aagcggccaa gcgaaagcgu uaaguauucc accugggga g uacgccggca 900  
 acggugaaac ucaaaggaaugacgggggc cngcacaagc ggaggaaacu gugguuuaau 960  
 ucgaugauac gcgaggaaacc uuaccggggc uuaaaauugca gaugaauua c ggugaaagcc 1020  
 guaagccgca agggcaucugu gaaggugcug caugguuguc gucagcucgu gccgugagggu 1080  
 gucggcuuaa gugccauaac gagcgcaacc cuuguuguca guuacuaac a gguuaugcug 1140  
 aggacucuga caagacugcc aucguaagau gugaggaagg uggggauga c gucaaaucag 1200  
 cacngcccuu acguccgggg cuacacacgu guuacaauagg gggguacaga gggcngcuac 1260  
 cacgcgagug gaugccaauc cccaaaaccu cucucaguuc ggacuggagru cugcaacccg 1320  
 acuccacgaa gcuggauucg cuaguaaucg cgcaucagcc acggcgcggu gaauacguuc 1380  
 ccgggccuug uacacaccgc ccgucaaguc augggagccg gggguaccug aagugcguaa 1440  
 ccgcgaggag cgcccuaggg uaaaacuggu gacuggggcu aagucguaa c aagguagcng 1500  
 uaccggaagg aacaccuccu uucu 1524

<210> 5

<211> 1173

5 <212> ARN

<213> Propionibacterium acnes

<400> 5

uuggagaguu ugauccuggc ucaggacgaa cgcugggcggc gugcuuaaca caugcaaguc 60  
 gaacggaaag gccucgcuuu uguggggugc ucgagugcg aacggguga g uaacacguga 120  
 guaaccugcc cuugacuug ggauaacuuc aggaaacugg ggcuaauac c ggauaggagc 180  
 uccugcugca uggugggggu uggaaaguuu cggcgguugg ggauggaalc c gggcuuauc 240  
 agcuuugugg uggguuagug gcuuaccaag gcuuugacgg guagccggc c ugagagggug 300  
 accggccaca uugggacuga gauacggccc agacuccuac gggaggcagc aguggggaa 360  
 auugcacaau gggcggaagc cugaugcagc aacgcccgcu gcggggauga c ggccuucggg 420  
 uuguaaaccc cuuucgccug ugacgaagcg ugagugacgg uauuggguaa agaagcacc 480  
 gcuaacuacg ugccagcagc cgcggugaua cguagggugc caacguuguc cggauuuuu 540  
 gggcguaaag ggcucguagg uggugaucg cgcgggaagu gaaucuuug ggcuaaaccc 600  
 ugagcgugcu uucgauacgg guugacuuga ggaagguagg ggagaugga auuccuggug 660  
 gagcguggga augcgagau aucaggagga acaccagug cgaaggcggu ucucugggcc 720  
 10 uuuccugacg cugaggagcg aaagcguggg gagcgaacag gcuuagauac ccugguaguc 780  
 cacgcuguaa acggugggua cuaggugugg gguccauucc acggguuccg ugccguagcu 840  
 aacgcuuaa guacccgcc uggggagugac ggcgcaagg cuaaaacuca aaggaauga 900  
 cggggcccc cacaagcggc ggagcaugc gauuaaucg augcaacgcg uagaaccuaa 960  
 ccuggguug acauggaucg ggagugcuca gagaugggug ugccucuuuu ggggucgguu 1020  
 cacagguggu gcauggcugu cguacgucg ugucgugaga uguugggua agucccgcaa 1080  
 cgagcgcaac ccuugucuc uguugccagc acguuaggu ggggacucag uggagaccgc 1140  
 cggggucaac ucggaggaag ggggggauga cgu 1173

<210> 6

15 <211> 1517

ES 2 391 744 T3

<212> ARN  
<213> Clostridium nexile

<220>  
5 <221> modified\_base  
<222> (184)..(185)  
<223> n es a, c, g, o u

10 <220>  
<221> modified\_base  
<222> (341)..(341)  
<223> n es a, c, g, o u

15 <220>  
<221> modified\_base  
<222> (430)..(430)  
<223> n es a, c, g, o u

20 <400> 6

```

gagauuugau ccuggcucag gaugaacgcu ggccggccgu gcuuacacau gcagucgaac 60
gaagcgcuaa aacuggauuu cuucggauug aaguuuuugc ugacugagug gcggacgggu 120
gaguaacgcg uggguaaccu gccucauaca gggggauaac aguuagaaau gcugcuaau 180
accnnauaag cgcacagugc ugcauggcac aguguaaaaa cuccgguggu augagaugga 240
cccgcgucug auuagcuagu ugguggggua acggccuacc aaggcgacga ucaguagccg 300
gccugagagg gugaacggcc acauugggac ugagacacgg nccaaacucc uacggggaggc 360
agcagugggg aaauuugcac aaugggggaa acccugaugc agcgacgccg cgugagcgaa 420
gaaguauuun gguauguaaa gcucuauacag cagggagaa aaugacggua ccugacuaag 480
aagcaccggc uaaauacgug ccagcagccg cgguaauacg uaggugcaag cguaucggg 540
auuuacuggg uguaaaggga gcguagacgg uuguguaagu cugaugugaa agcccggggc 600
ucaaccccgg acugcauugg aaacuaugua acuagagugu cggagaggua agcggaaauc 660
cuaguguagc ggugaauugc guagauuuu ggaggaacac caguggcgaa ggcggcuuac 720
uggacgauca cugacguuga ggucgaaag cguggggagc aaacaggauu agauaccug 780
guaguccacg ccguaaacga ugacuacuag gugcggggga gcaaagcucu ucggugccgc 840
agcaaacgca auaaguaguc caccuggggga guacguucgc aaga zugaaa cucaaaggaa 900
uugacgggga cccgcacaag cguggagcau gugguuuauu ucga gcaacg cgaagaccuu 960
accuggucuu gacaucggg ugaccggucc aguaauggga ccuuuccuuc gggacacggg 1020
gacaggguggu gcaugguugu cgucagcucg ugucgugaga uguuggguua agucccgcaa 1080
cgagcgcaac ccuauucuc aguagccagc auuuaggug ggca cucugg agagacugcc 1140
aggguaaacc uggaggagg uggggauagc gucaaaucuu caug ccccuu augaccaggg 1200
cuacacacgu gcuacaauug cguaaacaaa gggagcggaa ccugugaggg gaagcaaauc 1260
ucaaaaauaa cgucucaguu cggauuguag ucugcaacuc gacuacauga agcuggaauc 1320
gcuaguaauc gcgaucagc augucgcggg gaauacguuc cccgggucuu uacacaccgc 1380
ccgucacacc augggaguca guaacgcccg aagucaguga ccca accgua aggaggaggc 1440
ugccgaaggu gggaccgaa acugggguga agucguaaca aggu agccgu aucggaaggu 1500
gcggcuggau caccucc 1517

```

<210> 7  
 <211> 1480  
 <212> ARN  
 <213> Eubacterium formicigenerans  
 5  
 <220>  
 <221> modified\_base  
 <222> (1) .. (1)  
 <223> n es a, c, g, o u  
 10  
 <220>  
 <221> modified\_base  
 <222> (124) .. (126)  
 <223> n es a, c, g, o u  
 15  
 <220>  
 <221> modified\_base  
 <222> (221)..(223)  
 <223> n es a, c, g, o u  
 20  
 <220>  
 <221> modified\_base  
 <222> (288)..(288)  
 <223> n es a, c, g, o u  
 25  
 <220>  
 <221> modified\_base  
 <222> (350)..(351)  
 <223> n es a, c, g, o u  
 30  
 <220>  
 <221> modified\_base  
 <222> (518)..(518)  
 <223> n es a, c, g, o u  
 35  
 <220>  
 <221> modified\_base  
 <222> (535) .. (536)  
 <223> n es a, c, g, o u  
 40  
 <220>  
 <221> modified\_base  
 <222> (726)..(726)  
 <223> n es a, c, g, o u  
 45  
 <220>  
 <221> modified\_base  
 <222> (919)..(919)  
 <223> n es a, c, g, o u  
 50  
 <220>  
 <221> modified\_base  
 <222> (922)..(922)  
 <223> n es a, c, g, o u  
 55  
 <220>  
 <221> modified\_base  
 <222> (956)..(957)  
 <223> n es a, c, g, o u  
 60  
 <220>  
 <221> modified\_base  
 <222> (1185)..(1186)  
 <223> n es a, c, g, o u  
 65  
 <400> 7

ES 2 391 744 T3

nuuaaacgag aguuugaucc uggcucagga ugaacgcugg cggcgugcuu aacacaugca 60  
 agucgagcga agcacauaag uuugauucuu cggaugaaga cuuuugugac ugagcggcgg 120  
 acgnnnngagu aacgcguggg uaaccugccu cauacagggg gauaacagyu agaaauggcu 180  
 gcuaauaccg cauaagacca caguacugca ugguacagug nnnaaacuc cggugguaug 240  
 agauggacc cgcucugauu agguaguugg ugagguaacg gccaccnag cgcacgauca 300  
 guagccgacc ugagagggug accggccaca uugggacuga gacacggccn ngacuccuac 360  
 gggaggcagc aguggggaau auugcacaau gggcgaaagc cugaugcagc gacgccgcu 420  
 gaaggaugaa guauuucggg auguaaacuu cuaucagcag ggaagaaaau gacgguaccu 480  
 gacuaagaag ccccgccuaa cuacgugcca gacgccngg uauuacguag ggggnnagcg 540  
 uuauccggau uuacugggug uaaagggagc guagaccgcu gugcaagucu gaagugaaag 600  
 gcaugggcuc aaccugugga cugcuuugga aacugugcag cuagaguguc ggagagguaa 660  
 guggaaaucc uaguguagcg gugaauugcg uagauuuuag gaggaacacc aguggcgaag 720  
 gcggnucacu ggacgaugac ugacguugag gcucgaaaagc guggggagca aacaggauua 780  
 gauaccucgg uaguccacgc cguaaacgau gacugcuaagc ugucggguag caaagcuauu 840  
 cggugccgca gcuaacgcaa uaagcagucc accuggggag uacguucgca agaaugaaac 900  
 ucaaaggaa ugcaggggnc cngcacaagc gguggagc au gugguuuau ucgaannaac 960  
 gcgaagaacc uuaccugauc uugacaucucc gaugaccgcu ucguaaugga agyuuuucuu 1020  
 cggaaacucg gugacaggug gugcaugguu gucgucagcu cgugucguga gauguugggu 1080  
 uaagucccg aacgagcgca acccuuaucu ucaguaagca gcauuuagga ugggcaucuu 1140  
 ggagagacug ccagggaaua ccuggaggaa gguggggaug acgunnaauc aucaugcccc 1200  
 uuauagaccg ggcuaacacac gugcuacaau ggcguaaa ca gagggaggca gagccgcgag 1260  
 gccgagcaaa ucuaaaaau aacgucucag uucggauu gu agucugcaac ucgacuacau 1320  
 gaagcuggaa ucgcuaguaa ucgcagauca gaaugcugcg gugaauacgu ucccggguu 1380  
 uguacacacc gcccgucaca ccaugggagu caguaacg cc cgaagucagu gacccaaccg 1440  
 aaaggagga gcugccgaag gugggaccga uaacuggggu 1480

5

<210> 8  
 <211> 1423  
 <212> DNA  
 <213> Ruminococcus gnavus

10

<220>  
 <221> modified\_base  
 <222> (1148)..(1148)  
 <223> n e s a , c , g , o t

15

<220>  
 <221> modified\_base  
 <222> (1188)..(1188)  
 <223> n e s a , c , g , o t

20

<220>  
 <221> modified\_base  
 <222> (1257)..(1257)  
 <223> n e s a , c , g , o t

25

<400> 8

ES 2 391 744 T3

```

cctggctcag gatgaacgct ggcggcgtgc ttaacacatg caagtcgagc gaagcacctt    60
gacggatttc ttcggattga agccttgggtg actgagcggc ggaecgggtga gtaacgcgtg    120
ggtaacctgc ctcatacagg gggataacag ttggaaacgg ctgctaatac cgcataagcg    180
cacagtaccg catggtacgg tgtgaaaaac tccgggtgta tgagatggac ccgcgtctga    240
ttaggtagtt ggtggggtaa cggcctacca agccgacgat cagtagccga cctgagaggg    300
tgaccggcca cattgggact gagacacggc ccaaactcct acgggaggca gcagtgggga    360
atattgcaca atgggggaaa ccctgatgca gcgacgccgc gtgagcgatg aagtatttcg    420
gtatgtaaag ctctatcagc agggaagaaa atgacggtac ctgactaaga agccccggct    480
aactacgtgc cagcagccgc ggtaatacgt agggggcaag cgttatccgg atttactggg    540
tgtaaagggg gcgtagacgg catggcaagc cagatgtgaa agcccggggc tcaacccccg    600
gactgcattt ggaactgtca ggctagagtg tccgagagga aagcgggaatt cctagtgtag    660
cggtgaaatg cgtagatatt aggaggaaca ccagtggcga aggcggcttt ctggacgatg    720
actgacgttg aggctcgaaa gcgtggggag caaacaggat tagataccct ggtagtccac    780
gccgtaaacg atgaatacta ggtgtcgggt ggaaaagcca ttccggtgccg cagcaaacgc    840
aataagtatt ccacctgggg agtacgttcg caagaatgaa actcaaagga attgacgggg    900
accgcacaaa gcgggtggagc atgtggttta attcgaagca accgcaagaa ccttacctgg    960
tcttgacatc cctctgaccg ctctttaatc ggagctttcc ttccgggacag aggagacagg   1020
tgggtcatgg ttgtcgtcag ctcgtgtcgt gagatgttgg gttaagtccc gcaacgagcg   1080
caacccttat ctttagtagc cagcattttg gatgggcact ctagagagac tvccagggat   1140
aacctggngg aagggtggga tgacgtcaaa tcatcatgcc ccttatgncc agggetacac   1200
acgtgctaca atggcgtaaa caaaggggaag cgagcccgcg aggggggagca aatcccnaaa   1260
ataacgtctc agttcggatt gtagtctgca actcgactac atgaagctgg aategctagt   1320
aatcgegaat cagaatgtcg cgggtaatac gttcccgggt ctgtacaca ccseccgtca   1380
caccatggga gtmagtaacg cccgaagtca gtgacccaac cgc                                1423

```

5

```

<210> 9
<211> 1418
<212> DNA
<213> Ruminococcus torques

```

10

```

<220>
<221> modified_base
<222> (724)..(724)
<223> n es a, c, g, o t

```

15

```

<220>
<221> modified_base
<222> (771)..(771)
<223> n es a, c, g, o t

```

20

```

<220>
<221> modified_base
<222> (776)..(776)
<223> n es a, c, g, o t

```

25

```

<220>
<221> modified_base
<222> (807)..(807)

```

ES 2 391 744 T3

<223> n e s a, c, g, o t

<400> 9

	ctcaggatga acgctggcgg cgtgcctaac acatgcaagt cgagcgaagc actttgctta	60
	gattcttcgg atgaagagga ttgtgactga gcggcggacg ggtgagtaac gcgtgggtaa	120
	cctgcctcat acagggggat aacagttaga aatgactgct aataccgcat aagaccacag	180
	caccgcatgg tgcgggggta aaaactccgg tggatagaga tggaccgcg tctgattagc	240
	tagttggtaa ggtaacggct taccaaggcg acgatcagta gccgacctga gagggtgacc	300
	ggccacattg ggactgagac acggcccaaa ctctacggg aggcagcagt ggggaatatt	360
	gcacaatggg ggaaacctg atgcagcgac gccgcgtgag cgaagaagta tttcggtatg	420
	taaagctcta tcagcagggg aaaaaatgac ggtacctgac taagaagcac cggctaaata	480
	cgtgccagca gccgcggtaa tacgtatggt gcaagcgtta tccggattta ctgggtgtaa	540
	agggagcgtg gacggatggg caagtctgat gtgaaaacc ggggctcaac cccgggactg	600
	cattggaaac tgttcatcta gagtctgga gaggtaagtg gaattcctag ttagcgggtg	660
	aaatgcgtag atattaggag gaacaccagt ggcgaaggcg gcttactgga cagtaactga	720
	cgtngaggct cgaaagcgtg gggagcacac aggattagat accctggtag nccacnccgt	780
	aaacgatgac tactaggtgt cgggtgncaa agccattcgg tccgcagca aacgcaataa	840
	gtagtccacc tggggagtac gttcgcaaga atgaaactca aaggaattga cggggaccog	900
	cacaagcgtg ggagcatgtg gtttaattcg aagcaacgcg aagaacctta cctgctcttg	960
	acatcccgtc gaccggacgg taatgcgtcc ttcccttcgg ggcagcggag acaggtggtg	1020
	catggttgc gtcagctcgt gtcgtgagat gttgggttaa gtcccgcaac gagcgcgaacc	1080
	cctatcttta gtagccagcg gccaggccgg gcactctaga gagactgccg gggataaacc	1140
	ggaggaaggt ggggatgacg tcaaatcatc atgccctta tgagcagggc tacacacgtg	1200
5	ctacaatggc gtaaacaaag ggaagcgaga ccgcgaggtg gagcaaatcc caaaaataac	1260
	gtctcagttc ggattgtagt ctgcaactcg actacatgaa gctggaatcg ctagtatcg	1320
	cgaatcagaa tgtcgcggtg aatacgttcc cgggtcttgt acacaccgcc cgtcacacca	1380
	tgggagtcag taacgccga agtcagtgac ccaaccgt	1418

<210> 10

10 <211> 1458

<212> ARN

<213> Clostridium sp.

<400> 10

15

ES 2 391 744 T3

gaugaacgcu ggcggcgugc uuaacacaug caagucgagc gaagcgauuc uaawgaaguu 60  
 uucggaygga auuuraauug acugagcggc ggacggguga gaaacgcgug gguaaccugc 120  
 cucauacagg gggauaacag uuggaaaacgg cugcuaauac cgcauaagca cacagugccg 180  
 caugguacgg ugugaaaaac uccgguggua ugagauggac ccgcgucuga uuagguaguu 240  
 ggugagguaa cggcccacca agccgacgau caguagccga ccugagaggg ugaccggcca 300  
 cauugggacu gagacacggc ccaaaccuccu acgggaggca gcagugggga auauuggaca 360  
 augggggaaa ccugaucca gcgacccgc gugagugaag aaguauuucg guauguaaag 420  
 cucuaucagc agggaaagaaa augacggua cugacuaaga agccccggcu aacuaacgugc 480  
 cagcagccgc gguaauacgu agggggcaag cguuauccgg auuuacuggg uguaaagggg 540  
 gcguagacgg caaugcaagu cuggagugaa agccccggc ucaaccccgg gacugcuuug 600  
 gaaacugugu ugcuaagagug caggagaggu aaguggaauu ccuaguguag cggugaaauug 660  
 cguagauuu agggaggaaca ccaguggcga aggcggcuua cuggacugua acugacguug 720  
 aggcucgaaa gcguggggag caaacaggau uagauaccu gguaaguccac gccguaaacg 780  
 augaaucua gguguugggg agcaaagcuc uucggugccg ccgcuaacgc aauaaguauu 840  
 ccaccugggg aguacguucg caagaauagaa acuaaagga auugacgggg acccgcaaca 900  
 gcgugggagc augugguuuu auucgaagca acggaagaa ccuuaccaag ucuugacauc 960  
 ggaauagacc guccguaacg gggccuuccc uacggggcau uccagacagg uggugcaugg 1020  
 uugucgucag cucgugucgu gagauguuug guuaagucc gcaacgagcg caaccuuau 1080  
 ccuaguagc cagcauguag uggugggac ucuggggaga cugccaggga uaaccuggag 1140  
 gaaggugggg augacgucaa aucaucaugc cccuuauagau uggggcuaca cacgugcuac 1200  
 aauggcguaa acaaagggaa gcaaaggagc gaucuuuagc aaaccccaaa aauaacgucu 1260  
 caguucggau uguagucugc aacucgacua caugaagcug gaaucgcuag uaaucgcgga 1320  
 ucagaaugcc gcgguuaua cguucccggg ucuuguacac accgcccguc acaccauggg 1380  
 aguugguaac gcccgaguc agucgacccam ccsmaggag ggagcugccg aaggcgggac 1440  
 urauaacugg ggugaaug 1458

5 <210> 11  
 <211> 728  
 <212> ARN  
 <213> Moraxella cuniculi

10 <400> 11

ES 2 391 744 T3

aggcuuaaca caugcaaguc gaa cgaaguu agggagcuug cuccugauac uuaguggcgg 60  
 acgggugagru aaugcuuagg aau cugccua guaguggggg auaacuaucc gaaaggauag 120  
 cuaauaccgc auacgaccua cggguguaag ggggcuuaag cucucgcuau uagaugagcc 180  
 uaagucggau uagcuaguug gugggguaaa ggccuaccaa ggcgacgaur uguagcuggu 240  
 cugagaggau gaucagccac acu ggggacug agacacggcc cagacuccua cgggaggcag 300  
 caguggggaa uauuggacaa ugfgcgaaag ccugauccag ccaugcccgc gugugugaag 360  
 aaggccuuuu gguuguaaag cacuuuaagu ggggaggaaa agcuauuagc uauuaccuau 420  
 uagcccugac guuaccaca gaa uaaagcac cggcuaacuc ugugccagca gccgcgguaa 480  
 uacagagggu gcaagcguaa auc ggaauua cugggcuuaa agcgcgcuua gguuguuacu 540  
 uaagucagau gugaaagccc cgggcuuaac cugggaacug caucugauac uggguaacua 600  
 gaguagguga gagggaaqua gaa uuaccagg uguagcggug aaaugcguag agaucuggag 660  
 gaauaccgau ggcgaaggca gcu uccuggc aucauacuga cacugaggug cgaaagcgug 720  
 gguagcaa 728

<210> 12  
 <211> 1519  
 <212> ARN  
 <213> Moraxella lacunata

5

<400> 12

agaguugau cauggcucag auu gaaacgcu ggcggcaggc uuaacacaug caagucgaac 60  
 gaugaagucu agcuugcuag acggauuagu ggcgaacggg ugaguaaugc uuaggaucu 120  
 gccuauuagu gggggauaac guagggaaac uuacgcuaau accgcuaacg ucuuacgaga 180  
 gaaagggggc uuuuagcucu cgcuaauaga ugagccuag ucggauuagc uaguuggugg 240  
 gguaaaggcc uaccaaggcg acg aucugua gcuggucuga gaggaugauc agccacacug 300  
 ggacugagac acggccaga cuccuacggg aggcagcagu ggggaauuu ggacaauggg 360  
 cgaaagccug auccagccau gccgcgugug ugaagaaggc cuuuugguug uaaagcacuu 420

10



ES 2 391 744 T3

uaagugggga ggaaaagcuu gugguuaua ccuacaagcc cugacguuac ccacagaaua 480  
 agcaccggcu aacucugugc cagcagccgc gguaauacag agggugcaag cguuaaucgg 540  
 aauuacuggg cguaaagcga gcguaggugg ucuuuaagu cagaugugaa agccccgggc 600  
 uuaaccuggg aacugcaucu gauacugggu gacuagagua ggugagaggg aaguagaauu 660  
 ccaggugugag cggugaaaug cguagagau c uggaggaaua ccgauggcga aggcagcuuc 720  
 cuggcaucau acugacacug agguucgaa a gcguggguag caaacaggau uagauacccu 780  
 gguaguccac gccguaaacg augucuacc a gucguugggu cucuugaaga cuuagugacg 840  
 caguuaacgc aauaaguaga ccgccugggg aguacggccg caagguaaaa acucaaauga 900  
 auugacgggg cccgcacaag cgguggagca ugugguuuaa uucgaugcaa cgcgaagaac 960  
 cuuaccuggu cuugacauag ugagaauccu gcagagaugc gggagugcuu cgggaaauca 1020  
 cauacaggug cugcauggcu gucgucagcc cgugucguga gauguugggu uaaguccgc 1080  
 aacgagcgca acccuuuucc uuaguuaacc a gcgauuaag ucgggaacuc uaaggauacu 1140  
 gccagugaca aacuggagga aggcgggacg acgucaaguc aucauggccc uuacgaccag 1200  
 ggcuaacacac gugcuacaau gguugguacc a aaggguuugcu acacagcgau gugaugcuaa 1260  
 ucucaaaaag ccaaucguag uccggauugg agucugcaac ucgacuccau gaagucggaa 1320  
 ucgcuaguaa ucgcagauca gaaugcugcg gugaauacgu ucccgggccu uguacacacc 1380  
 gcccucacaca ccaugggagu ugauucaccc agaagugguu agccuaacgc aagagggcga 1440  
 ucaccacggg ggggucgag acuggggugc agucguaa ca agguagccgu aggggaaucg 1500  
 cggugggauc accucuuu 1519

<210> 13  
 <211> 1448  
 <212> ARN  
 <213> Moraxella osloensis

5

<400> 13

cuggcggcag gcuuaacaca ugcaagucga acgaugacuc ucuagcuugc uagagaugau 60  
 uaguggcggg cgggugagua acuuuagga aucugccuag uaguggggga uagcucgggg 120  
 aaacucgaau uaauaccgca uacgaccuac gggugaaagg gggcgcaagc ucuugcuauu 180  
 agaugagccu aaucagauu agcuaguugg ugggguaaag gccaccacag gcgacgaucu 240  
 guaacugguc ugagaggau gacagucaca ccggaacuga gacacggucc ggacuccuac 300  
 gggagggcagc aguggggaau auuggacaau gggggcaacc cugauccagc caugcccgcu 360

10

ES 2 391 744 T3

gugugaagaa ggccuuuugg uuguaaagca cuuuuagcag ggaggagagg cuaaugguua 420  
 auacccauua gauuagacgu uaccuagcaga auuagcaccg gcuaacucug ugccagcagc 480  
 cgcgguuaa cagagggucc gagcguuaau cggaauuacu gggcguaaag cgaguguagg 540  
 uggcucauuu agucacaugu gaaaaccucc ggcuuaaccu ggggaacugca ugugauacug 600  
 guugugcuag aauaugugag agggaaaguag aauuccaggu guagcgguga aaugcguaga 660  
 gaucuggagg aauaccgaug gcgaaggcag cuuccuggca uauauugac acugagauuc 720  
 gaaagcguug guagcaaca ggauuagaua cccugguagu ccacgccgua aacgaugucu 780  
 acuagccguu gggguccuug agacuuuagu ggcgcaguua acgcgaaag uagaccgccu 840  
 ggggaguacg gccgcaaggu uaaaacucua augaauugac gggggcccgc acaagcggug 900  
 gagcaugugg uuuauucga ugcaaccgca agaaccuac cuggucuuga cauagugaga 960  
 aucucucaga gaugagagag ugcclucggg aacucacua caggugcugc auggcugucg 1020  
 ucagcucgug ucgugagaug ugggguuaag ucccgcaccg agcgcacc cc uuuuccuuu 1080  
 uugccagcgg guuaagccgg gaacuuuaag guacugcca gugacaaacu ggaggaggc 1140  
 ggggacgacg ucaagcauc auggcuccua cgaccagggc uacacacgug cuacauggu 1200  
 agguacagag gguugcuaca cagcgaugug augcuauuc caaaaagccu aucguagucc 1260  
 ggauuggagu cugcaacucg acucCaugaa gucggaaucg cuaguaaucg cggaucagaa 1320  
 ugcgcggug aauacgucc cgggCcuugu acacaccgcc cgucacacca ugggagucua 1380  
 uugcaccaga aguagguagc cuaaCgaaag agggcgcuua ccacggugug gucgaugacu 1440  
 ggggugaa 1448

<210> 14  
 <211> 452  
 <212> ARN  
 <213> Acinetobacter junii

5

<400> 14

caugcaaguc gagcggagau gaggugcuug caccuuuacu uagcggcgga cgggugagua 60  
 augcuuagga aucugccuau uaguaggggga caacauuccg aaaggaaugc uauuaccgca 120  
 uacguccuac ggggagaaagc agggggaucuu cggaccuugc gcuaauagau gagccuaagu 180  
 cggauuagcu aguugguggg guaaaggccu accaaggcga cgauucugag cgggucugag 240  
 aggaugaucc gccacacugg gacugagaca cggcccagac uccuacggga ggcagcagug 300  
 gggaaauuug gacaaugggg ggaaCccuga uccagccaug ccgcgugugu gaagaaggcc 360  
 uuaugguugu aaagcacuuu aagcGaggag gaggcuacug agacuaauac ucuuggauag 420  
 uggacguuac ugcgagaaua agcaccggcu aa 452

10

<210> 15  
 <211> 1488  
 <212> ARN  
 <213> Comamonas sp.

15

<400> 15

20

ES 2 391 744 T3

auugaacgcu ggcggcaugc cuuacacaug caagucgaac gguaacaggu cuucggaugc 60  
 ugacgagugg cgaacgggug aguaauacau cggaacgugc ccgauygggg gggauaacga 120  
 ggcgaaagcu uugcuauaac cgcauacgau cuacggauga aagcggggga ucuuaggacc 180  
 ucgcgaggac ggagcggccg auggcagauu agguaguugg ugggauaaaa gcuaa ccaag 240  
 ccgacgaucg guagcugguc ugagaggaug aucagccaca cugggacuga gacacggccc 300  
 agacuccuac gggaggcagc aguggggaa uuuaggacaau gggggaaacc cugauccagc 360  
 caugcccgcu gcagggaugaa ggccuucggg uuguaaacug cuuuugucg gaacgaaaag 420  
 gucucuucua auaaaggggg ccgaugacgg uaccguaaga auaagcaccg gcuaa cuacg 480  
 ugccagcagc cgcgguaaua cguagggugc aagcguuaau cggaaauacu gggcguaaag 540  
 cgugcgcagg cgguaaugua agacagauu gaaauccccg ggcucaaccu gggaa.cugca 600  
 uuugugacug cauggcuuga gugcggcaga gggggauugg auuccgugug uagca.gugaa 660  
 augcguagau augcggagga acaccgaugg cgaaggcaau ccccugggcc ugcaaugacg 720  
 cucaugcagc aaagcugggg gagcaaacag gauuagauac ccugguaguc cacgcccuaa 780  
 acgauguaa cugguugug ggaauuuguu uucucaguaa cgaagcuaac gcgugraagu 840  
 gaccgccugg ggaguacggc cgaagguug aaacuaaag gaauugacgg ggaccgcgac 900  
 aagcggugga ugaugugguu uaauucgag caacgcgaaa aaccuuacc accuuugaca 960  
 uggcaggaag uccacagaga ugaggaugug cucgaaagag aaccugcaca caggugcugc 1020  
 auggcugugc ucagcugug ucgugagaug uuggguuaag ucccgcaacg agcgaaccc 1080  
 uugucauuag uugcuacau uaguugggca cucuaaugag acugccggug acaaacggga 1140  
 ggaagguggg gaugacguca aguccuauug gccuuauag guggggcuac acacgucuaa 1200  
 caauggcugc uacaaagggu ugccaacccg cgaaggggag cuaaucccua aaagccaguc 1260  
 guaguccgga ucgcagucug caacucgacu gcgugaaguc ggaaucgua gaaacgugg 1320  
 aucagaauug cacggugaau acguucccg gucuuguaca caccgcccg cacaccaugg 1380  
 gagcgggucu cgcagaagu agguagccua accgcaagga gggcgcuuac cacgagggg 1440  
 uucgugacug gggugaaguc gaaacaaggu agccguaucg gaaggugc 1488

5 <210> 16  
 <211> 1520  
 <212> ARN  
 <213> Comamonas terrigena

10 <220>  
 <221> modified\_base  
 <222> (623)..(623)  
 <223> n es a, c, g, o u

15 <220>  
 <221> modified\_base  
 <222> (630)..(630)  
 <223> n es a, c, g, o u

20 <400> 16

ES 2 391 744 T3

aguuugaucc uggcucagau ugaacgcugg cggcaugcuu uacacaugca agucgaacgg 60  
 cagcacggac uucggucugg uggcgagugg cgaacgggug aguaauacau cggaacgu gc 120  
 ccaguugugg gggauaacua cucgaaagag uagcuaauac cgcaugagua cugagguu ga 180  
 aagcagggga ucgcaagacc uugcgcaacu ggagcggccg auggcagauu agguaguu gg 240  
 ugggauaaaa gcuuaccaag ccgacgaucu guagcugguc ugagaggacg accagcca ca 300  
 cugggacuga gacacggccc agacuccuac gggaggcagc aguggggaau uuuggaca au 360  
 gggcgaaagc cugauccagc aaugccgcgu gcaggaukaa ggccuucggg uuguaaacug 420  
 cuuuuguaag gaacgaaaag cuucggguua auaccuggag ucaugacggg acccuaagaa 480  
 uaagcaccgg cuaacuacgu gccagcagcc gcguaauac guagggugca agcguuaa uc 540  
 ggaauuacug ggcguaaagc gugcgagggc ggucuuuaa gacagaggug aaaucccgg 600  
 gcucaaccug ggaacugccu uunugacugn aaggcuggag ugcggcagag ggggauggaa 660  
 uuccgcgugu agcagugaaa ugcguagaua ugcggaggaa caccgauggc gaaggcaa uc 720  
 cccugggccu gcacugacgc ucaugcacga aagcggggg agcaaacagg auuagaua cc 780  
 cugguagucc acggccuaaa cgauguaaac ugguuuuugg gucuuuacug acucagua ac 840  
 gaagcuaacg cgugaaguug accgccuggg gaguacggcc gcaagguuga aacucaaa gg 900  
 aaauagcggg gacccgcaca agcgguggau gaugugguuu aaauugaugc aacgcgaa aa 960  
 accuuaccca ccuuugacau guacggaau cuuuagagau agaggagugc ucgaaaga ga 1020  
 gccguaacac aggugcugca uggcugucgu cagcucgugu cgugagaugu uggguuaa gu 1080  
 cccgcaacga gcgcaaccu ugccaauagu ugcuaagaaa gggcacucua augggacu uc 1140  
 cggugacaaa ccggaggaag guggggaua gcuaagucc ucauggccu uauagguggg 1200  
 gcuacacacg ucauacaauug gcugguacaa agggugucca acccgcgagg gggagcuauu 1260  
 cccauaaagc cagucguagu ccgguucgca gucugcaacu cgacugcgug aagucggaau 1320  
 cgcuaguuuu cguggaucag aaugucacgg ugaauacguu cccgggucuu guacacaccg 1380  
 cccguacac caugggagcg ggucucgcca gaaguaggua gccuaaccgu aaggaggggc 1440  
 cuuaccacgg cgggguuucgu gacuggggug aagucguaac aagguagccg uaucggaagg 1500  
 ugcggcugga ucaccuccu 1520

5 <210> 17  
 <211> 1476  
 <212> ARN  
 <213> Bergeyella zoohelcum

10 <220>  
 <221> modified\_base  
 <222> (138)..(138)  
 <223> n es a, c, g, o u

15 <220>  
 <221> modified\_base  
 <222> (141) .. (143)  
 <223> n es a, c, g, o u

20 <220>  
 <221> modified\_base  
 <222> (183)..(183)  
 <223> n es a, c, g, o u

<220>  
 <221> modified\_base  
 <222> (278) .. (278)  
 <223> n es a, c, g, o u  
 5  
 <220>  
 <221> modified\_base  
 <222> (285) .. (285)  
 <223> n es a, c, g, o u  
 10  
 <220>  
 <221> modified\_base  
 <222> (317)..(318)  
 <223> n es a, c, g, o u  
 15  
 <220>  
 <221> modified\_base  
 <222> (341) .. (342)  
 <223> n es a, c, g, o u  
 20  
 <220>  
 <221> modified\_base  
 <222> (344)..(344)  
 <223> n es a, c, g, o u  
 25  
 <220>  
 <221> modified\_base  
 <222> (528) .. (528)  
 <223> n es a, c, g, o u  
 30  
 <220>  
 <221> modified\_base  
 <222> (532)..(533)  
 <223> n es a, c, g, o u  
 35  
 <220>  
 <221> modified\_base  
 <222> (542)..(542)  
 <223> n es a, c, g, o u  
 40  
 <220>  
 <221> modified\_base  
 <222> (547)..(547)  
 <223> n es a, c, g, o u  
 45  
 <220>  
 <221> modified\_base  
 <222> (737)..(737)  
 <223> n es a, c, g, o u  
 50  
 <220>  
 <221> modified\_base  
 <222> (766)..(767)  
 <223> n es a, c, g, o u  
 55  
 <220>  
 <221> modified\_base  
 <222> (965)..(966)  
 <223> n es a, c, g, o u  
 60  
 <220>  
 <221> modified\_base  
 <222> (1022)..(1022)  
 <223> n es a, c, g, o u  
 65  
 <220>

ES 2 391 744 T3

<221> modified\_base  
 <222> (1085)..(1086)  
 <223> n es a, c, g, o u

5 <220>  
 <221> modified\_base  
 <222> (1096)..(1096)  
 <223> n es a, c, g, o u

10 <220>  
 <221> modified\_base  
 <222> (1205)..(1205)  
 <223> n es a, c, g, o u

15 <220>  
 <221> modified\_base  
 <222> (1288)..(1289)  
 <223> n es a, c, g, o u

20 <220>  
 <221> modified\_base  
 <222> (1384)..(1384)  
 <223> n es a, c, g, o u

25 <220>  
 <221> modified\_base  
 <222> (1406)..(1406)  
 <223> n es a, c, g, o u

30 <400> 17

```

auacaaugga gaguuugauc cuggcucagc augaacgcua gcgggaggcc uaacacaugc      60
aagccgagcg ggauuuguug guuagcuugc uaacuaacaa ugagagcggc guacggguugc      120
guaacacgug ugcaaccngc nnnuauucgg gggauagccu uucgaaagga agauuaauac      180
ccnauaauau auugauuggc aucaguuauu auugaaagcu ccggcgggaa gagaugggca      240
cgcguaagau uagcuaguug gugagguaac ggcucacnaa ggcuncgaur uuuagggggc      300
cugagagggg guaccnnaac acugguacug agacacggac nngncuccua cgggagggcag      360
cagugaggaa uauuggacaa ugggugagag ccugauccag ccaucccgcg ugaaggacua      420
aggaccuauug guuuguaaac uucuuuuaua caggggaaaa ccuacucucg ugaggguaagc      480
ugaagguacu guaugaaaua gcaccggcua acuccgugcc agcagccncg gnnauaccga      540
gngugcnaag guuauccgga uuuauugggu uuaaaggguc cguaggcggg ucgauaaguc      600
aguggugaaa gccugcagcu uaacuguaa acugccguug auacugucgg ucugagugc      660
auuugaggua gcuggaaua guaguguagc ggugaaaugc auagauuuu cucagaaac      720
caauugcgaa ggcaggnuac caaguuacaa cugacgcuga uggacnnaag cguggggagc      780
gaacaggauu agauaccucg guaguccacg cuguaaacga ugcuaacucg uuuuuggggu      840
auuauacuuc agagaccaag cgaagugau aaguuagcca ccuggggagu acgaaacga      900
guuugaaacu caaaggaaau gacgggggccc cgcacaagcg guggauuugc ugguuuaau      960
cgaunnuacg cgaggaaccu uaccaagacu uaaaugggaa uugacagcug uagaaauac      1020
gnuuucucg gacaauuuuc aaggugcugc augguugucg ucagcucgug ccgugaggg      1080
uuagnnuagc uccugnaacg agcgcaaccc cugucacuag uugccaucu uaguuggg      1140
  
```

ES 2 391 744 T3

acucuaguga gacucgcccuaac gcaaguagag aggaaggugg ggauagcguac aaaucaucac 1200  
 ggccnuuacg ucuuaggcca cacacguaaau acaauggccg guacagaggg cagcuacacu 1260  
 gcgaagugau gcgaauucucg aaagccgnnc ucaguucgga uuggagucug caacucgacu 1320  
 cuaugaagcu ggaaucgcu gaaucgcgc aucagccaug gcgaggugaa uacguucccg 1380  
 ggenuugac acacagcccg ucaagncaug gaagucuggg guaccugaag ucggugaccg 1440  
 uuaaaggagc ugccuagggu aaaacaggua acuaagg 1476  
  
 <210> 18  
 <211> 395  
 5 <212> ARN  
 <213> Chryseobacterium balustinum str. SBR1044  
  
 <220>  
 <221> modified\_base  
 10 <222> (343) .. (343)  
 <223> n es a, c, g, o u  
  
 <400> 18  
  
 uagcgggagg gcua.acacau gcaagccgag cgguaauuguu ucuucggaaa ugagagagcgg 60  
 gcguacgggu gcggaacaag ugugcaaccu gccuuuauuc gggggauagc cuuucgaaaag 120  
 gaagauuuau acuc.cauaac auauugaacg gcaucguuug auauugaaag cuccggcgga 180  
 uagagauggg caccg.cgcaag auuagcuagu uggugaggua acggcucacc aaggcgauga 240  
 ucuuuagggg ggcu.gagagg gugauccccc acacugguac ugagacacgg accagacucc 300  
 uacggggaggc agca.gugagg aaauuugguc aaugggugca agncugaacc agccaucaccg 360  
 15 cgugaaggac gacu.gcccuu aggguuuguaa acuuu 395  
  
 <210> 19  
 <211> 395  
 <212> ARN  
 20 <213> Chryseobacterium balustinum str. SBR2024  
  
 <220>  
 <221> modified\_base  
 <222> (11) .. (11)  
 25 <223> n es a, c, g, o u  
  
 <220>  
 <221> modified\_base  
 <222> (14) .. (14)  
 30 <223> n es a, c, g, o u  
  
 <220>  
 <221> modified\_base  
 <222> (17)..(18)  
 35 <223> n es a, c, g, o u  
  
 <220>  
 <221> modified\_base  
 <222> (310)..(310)  
 40 <223> n es a, c, g, o u  
  
 <400> 19

ES 2 391 744 T3

uagcgggagg ncunacnnau gcaagccgag cgguauguu ucuucggaaa uga gagagcg 60  
 gcguaacgggu gcggaacacg ugugcaaccu gccuuuaucu gggggauagc cuuucgaaag 120  
 gaagauuuuu acuccauaac auauugauug gcaucauuua auauugaaag cuc cggcggga 180  
 uagagauggg caccgcgaag auuagcuagu uggugaggua acggcucacc aagggcgauga 240  
 ucuuuagggg ggcugagagg gugaucccc acacugguac ugagacacgg accagacucc 300  
 uacgggaggg agcagugagg aaauugguc aaugggugca agccugaacc agc caucccg 360  
 cguagaaggac gacugcccua uggguuguaa acuuu 395

<210> 20

<211> 1435

5 <212> ARN

<213> Clostridium paraputrificum

<400> 20

cgaacgcugg cggcugccu aacacaugca agucgagcga ugaaguuccu ucgggaaacgg 60  
 auuagcggcg gacgggugag uaacacgugg gcaaccugcc uuauagaggg gaauagccuu 120  
 ccgaaaggaa gauuaauacc gcauaagauu guagcuucgc augaaguagc aaauaaaagga 180  
 gcaauccgcu auaaagauagg cccgcggcgc auuagcuagu uggugaggua acggcucacc 240  
 aaggcgacga ucgguagccg accugagagg gugaucggcc acauugggac uga gacacgg 300  
 cccagacucc uacgggaggc agcagugggg aaauuugcac aaugggggaa acccugaugc 360  
 agcaacgccg cgugagugau gacggccuuc gggguuuuaa gcucugucuu uggggacgau 420  
 aaugacggua cccaaggagg aagccacggc uaacuacgug ccagcagccg cgguaauacg 480  
 uagguaggaa gcguuugccg gauuuacugg gcguaaaggg agcguaggcg galuuuuuag 540  
 ugggagugua aaauaccggg cucaaccugg gugcugcauu ccaaacugga aaucuaagagu 600  
 gcagaggggg aaaguggaau uccuagugua gcguguaaau gcguagagau uaaggagaac 660  
 accaguggcg aaggcgacuu ucuggacugu aacugacgcu gaggcucgaa agcguugggga 720  
 10 gcaaacaggga uuagauacc ugguaugucca cgccguaaac gaugaauacu agguugaggg 780  
 guugucauga ccucugugcc gccgcuaacg cauuuaguuu uccgccuggg gagua cgguc 840  
 gcaagauuaa aacucaaaag aaugacggg gggccgcaca aguagcggag caugugguuu 900  
 aaucggaagc aacgcgaaga accuuaaccua gacuugacau cuccugaauu accau guaau 960  
 gugggaaguc ccuucgggga caggaaagaca gggugggcau gguugucguc agcucguguc 1020  
 gugagauuuu ggguaaaguc ccgcaacgag cgcaaccuuu auuguuaguu gcuac cauuu 1080  
 aguugagcac ucuagcgaga cugcccgggu uaacggggag gaaggugggg augac gucaa 1140  
 aucaucaugc ccuuuauugc uagggcuaca cacgugcuac aauggccggg acaacgagau 1200  
 gcaauaccgu gagguggagc aaaacuauaa aaccggucuc agnuccggaau guaggcugaa 1260  
 acucgccuac augaagcugg aguuacuagu aaucgcgaa cagaauuguc cggugaauac 1320  
 guucccgggc cuuguaacaca ccgccgguca caccaugaga guuggcaaua cccaa aguuu 1380  
 gugaucuaac ccguaaggga ggaagccacc uaagguaggg ucagcgauug gggug 1435

<210> 21

<211> 1509

15 <212> ARN

<213> Enterococcus cecorum



ES 2 391 744 T3

<400> 21

gacgaacgcu ggcggcgugc cuaauacaug caagucgaac gcauuuucuu ucaccgugac 60  
uugcuacacc ggaagaaaau gaguggcgaa cgggugagua acacgugggu aaccugccca 120  
ucagcggggg auaacaauug gaaacaggug cuaauaccgc auaauuccau uuaccgcaug 180  
guagugggau gaaaggcgcu uuugcgucac ugauuggaug acccggggug cauuagcuag 240  
uuggugggggu aacggccuac caaggcugcg augcauagcc gaccugagag ggugaucggc 300  
cacacuggga cugagacacg gcccagacuc cuacgggagg cagcaguagg gaaucuucgg 360  
caauggacgc aagucugacc gagcaacgcc gcgugaguga agaagguuuu cggaucguaa 420  
aacucuguuu uuagagaaga acaaggaua gaguggaaag uucauccuu gacgguaucu 480  
aaccagaaag ccacggcuua cuacgugcca gcagccggcg uaaucguag guggcaagcg 540  
uuguccggau uuauuggcg uaaagcgagc gcagggcguc uuuuaagucu gauugaaaag 600  
ccccggcuu aaccggggag ggucauugga aacugggaga cuugagugca gaagaaggaa 660  
gcggaauucc augugugagc gugaauugcg uagauauaug gaggaaacc aguggcgaag 720  
gcgcuuuucu gguucuuaac ugacgucgag gcucgaaagc guggggagca aacaggauua 780  
gauaccugg uaguccacgc cguaaacgau gagugcuaag uguuggaggg uuuccggccu 840  
ucagugcugc agcaaacgca uuaagcacuc cgccugggga guacgaccgc aagguaaaaa 900  
cucaaaggaa uugacgggga cccgcacaag cgguggagca ugugguuuaa uucgaagcaa 960  
cgcaagaac cuuaccaggu cuugacauc uuugaccauc cuagagauag gauuuuccu 1020  
ucggggacaa agugacaggu ggugcauggu ugucgucagc ucgugucgug agauguugg 1080  
uaaagucccg caacgagcgc aaccuuuuu guuaguugcc aucauucagu ugggcacucu 1140  
agcgagacug ccgcagacaa ugccggaggaa gguggggaug acgucaaac aucaugccc 1200  
uuauagaccug ggcuaacacac gugcuacaau ggagaguaca acgagucgca aagccgagag 1260  
gcuaaagcaa ucucuuaaag cucuucucag uucggauugu aggcugcaac ucgccuacau 1320  
gaagccggaa ucgcuaguaa ucgcggauca gcacgcccg gugaauacgu ucccggu cu 1380  
uguacacacc gcccgucaca ccacgagagu uuguaacacc caaagccggu gcgguaaccg 1440  
caaggagcca gccgucuaag gugggauaga ugauuggggg gaagucguaa caagguagcc 1500  
guaucggaa 1509

5 <210> 22  
<211> 1493  
<212> ARN  
<213> Enterococcus columbae

10 <220>  
<221> modified\_base  
<222> (33)..(33)  
<223> n es a, c, g, o u

15 <220>  
<221> modified\_base  
<222> (103)..(103)  
<223> n es a, c, g, o u

20 <220>  
<221> modified\_base

<222> (158)..(158)  
 <223> n es a, c, g, o u  
  
 <220>  
 5 <221> modified\_base  
 <222> (194)..(196)  
 <223> n es a, c, g, o u  
  
 <220>  
 10 <221> modified\_base  
 <222> (209) .. (209)  
 <223> n es a, c, g, o u  
  
 <220>  
 15 <221> modified\_base  
 <222> (230)..(230)  
 <223> n es a, c, g, o u  
  
 <220>  
 20 <221> modified\_base  
 <222> (254)..(254)  
 <223> n es a, c, g, o u  
  
 <220>  
 25 <221> modified\_base  
 <222> (289)..(289)  
 <223> n es a, c, g, o u  
  
 <220>  
 30 <221> modified\_base  
 <222> (316)..(316)  
 <223> n es a, c, g, o u  
  
 <220>  
 35 <221> modified\_base  
 <222> (346)..(347)  
 <223> n es a, c, g, o u  
  
 <220>  
 40 <221> modified\_base  
 <222> (367)..(367)  
 <223> n es a, c, g, o u  
  
 <220>  
 45 <221> modified\_base  
 <222> (430)..(432)  
 <223> n es a, c, g, o u  
  
 <220>  
 50 <221> modified\_base  
 <222> (441) .. (441)  
 <223> n es a, c, g, o u  
  
 <220>  
 55 <221> modified\_base  
 <222> (492) .. (492)  
 <223> n es a, c, g, o u  
  
 <220>  
 60 <221> modified\_base  
 <222> (555)..(556)  
 <223> n es a, c, g, o u  
  
 <220>  
 65 <221> modified\_base  
 <222> (565)..(565)

<223> n es a, c, g, o u  
 <220>  
 <221> modified\_base  
 5 <222> (613)..(613)  
 <223> n es a, c, g, o u  
 <220>  
 <221> modified\_base  
 10 <222> (639)..(639)  
 <223> n es a, c, g, o u  
 <220>  
 <221> modified\_base  
 15 <222> (783)..(783)  
 <223> n es a, c, g, o u  
 <220>  
 <221> modified\_base  
 20 <222> (788)..(788)  
 <223> n es a, c, g, o u  
 <220>  
 <221> modified\_base  
 25 <222> (792)..(792)  
 <223> n es a, c, g, o u  
 <220>  
 <221> modified\_base  
 30 <222> (811) .. (811)  
 <223> n es a, c, g, o u  
 <220>  
 <221> modified\_base  
 35 <222> (871)..(871)  
 <223> n es a, c, g, o u  
 <220>  
 <221> modified\_base  
 40 <222> (898)..(898)  
 <223> n es a, c, g, o u  
 <220>  
 <221> modified\_base  
 45 <222> (965)..(965)  
 <223> n es a, c, g, o u  
 <220>  
 <221> modified\_base  
 50 <222> (979)..(979)  
 <223> n es a, c, g, o u  
 <220>  
 <221> modified\_base  
 55 <222> (1013)..(1013)  
 <223> n es a, c, g, o u  
 <220>  
 <221> modified\_base  
 60 <222> (1035)..(1037)  
 <223> n es a, c, g, o u  
 <220>  
 <221> modified\_base  
 65 <222> (1063)..(1063)  
 <223> n es a, c, g, o u

<220>  
 <221> modified\_base  
 <222> (1068)..(1068)  
 5 <223> n es a, c, g, o u

<220>  
 <221> modified\_base  
 <222> (1071)..(1071)  
 10 <223> n es a, c, g, o u

<220>  
 <221> modified\_base  
 <222> (1102)..(1102)  
 15 <223> n es a, c, g, o u

<220>  
 <221> modified\_base  
 <222> (1127) .. (1128)  
 20 <223> n es a, c, g, o u

<220>  
 <221> modified\_base  
 <222> (1228)..(1228)  
 25 <223> n es a, c, g, o u

<220>  
 <221> modified\_base  
 <222> (1337)..(1337)  
 30 <223> n es a, c, g, o u

<220>  
 <221> modified\_base  
 <222> (1399)..(1399)  
 35 <223> n es a, c, g, o u

<220>  
 <221> modified\_base  
 <222> (1401)..(1401)  
 40 <223> n es a, c, g, o u

<220>  
 <221> modified\_base  
 <222> (1424) .. (1424)  
 45 <223> n es a, c, g, o u

<220>  
 <221> modified\_base  
 <222> (1462)..(1462)  
 50 <223> n es a, c, g, o u

<400> 22

ES 2 391 744 T3

ugagaguuug auccuggcuc aggacgaaacg cungcggcgu gccuaauaca ugcaagucga 60  
 acgcacuuuc uuucaccgua gcuugcuaca ccgaaaguaa gunaguggcg aacgggugag 120  
 uuacacgugg guaaccugcc caucagcggg ggauaacncu uggaaacagg ugcuaauacc 180  
 gcauaauuuu acunnnngca ugagaaguna uugaaaggcg caacuycgun acugauggau 240  
 ggaccgcggg ugcnuuagcu aguuggugag gnaacggccu accaaagcna cgaugcauag 300  
 ccgaccugag agggunaucg gccacacugg gacugagaca cggccnnaac uccuacggga 360  
 ggcagcngua gggaaucuuu ggcaauggac gcaagucuga ccgagcaacg ccgcgugagu 420  
 gaagaaggun nncggauvcu naaacucugu uguuagagaa gaacaaggau gagaguggaa 480  
 aguucauccc ungacggauu cuaaccagaa agccacggcu aacuaagcgc cagcagccgc 540  
 gguaauacgu aggunncaag cguunuccgg auuuuuuggg cguaaagcga gcgcagggcg 600  
 ucuuuuaagu cunaugugaa agcccacggc uuaaccgung agggucuuug gaaacuggga 660  
 gacuugagug cagaagagga aagcggaaau ccauguguaq cggugaaaug cguagauua 720  
 ugcaggaaca ccaguggcga agggggcuuu cuggucugua acugaagcug aggcucgaaa 780  
 gcnuaggngag gnaacaggau uagauaccu nguaguccac gccguaaaag augagucua 840  
 aguguuggag gguuuccgcc cuucagugcu ncagcaaacg cauaaagcgc uccgccungg 900  
 gaguacgacc gcaagguuga aacucaaaag aauugacggg gaccgcacaa gcgguggagc 960  
 augunguuua auucgaagna acgcgaagaa ccuuaccagg ucuugacauc cunugaccau 1020  
 ccuagagaua ggactnnccu ucggggacaa agugacaggu gngc'aungu ngucgucagc 1080  
 ucgugucgug agauguuggg unaagucccg caacgagcgc aacc'cunnuu guuaguuvc 1140  
 aucauuuagu ugggcacucu agcgagacug ccgcagacaa ugcg'raggaa gguggggauq 1200  
 acgucaaauc aucaugcccc uuaugacnug ggcuaacacac gugcu'acaau ggagaguaca 1260  
 acgaguugcg aagucgugag gcuaaagcuaa ucucuuaaag cucu'ucucag uucggauugu 1320  
 aggcugcaac ucgccuncau gaagccggaa ugcuaaguaa ugcg'gauca gcacgcccgc 1380  
 gugaauacgu ucccgggnlu nguacacacc gcccgucaca ccangagagu uuguaacacc 1440  
 cgaagccggg gggguaacgg cnaggagcca gccgucuaag gugg'gauaga uga 1493

<210> 23

<211> 1535

5 <212> ARN

<213> Enterococcus hirae

<400> 23

ES 2 391 744 T3

ccuggcucag gacgaacgcu ggcggcgguc. cuaauacaug caagucgaac gcuucuuuuu 60  
ccaccggagc uugcuccacc ggaaaaagag gaguggcgaa cgggugagua acacgugggu 120  
aaccugccca ucagaagggg auaacacuug gaaacaggug cuaauaccgu auaacaaucg 180  
aaaccgcaug guuuugauuu gaaaggcgc uucggguguc gcugauggau ggaccgcgg 240  
ugcauuagcu aguuggugag guaacggcuc accaaggcga cgaugcauag ccgaccugag 300  
agggugaucg gccacauugg gacugagaca cggcccaaac uccuacggga ggcagcagua 360  
gggaaucuuu ggcaauggac gaaagucuga ccgagcaacg ccgcgugagu gaagaagguu 420  
uucggaucgu aaaacucugu uguuagagaa gaacaaggau gagaguaacu guucauccu 480  
ugacgguauc uaaccagaaa gccacggcua acuacgugcc agcagccgcg guaaucgua 540  
gguggcaagc guuguccgga uuuauuggc guaaagcgag cgcagggcgg uucuaaguc 600  
ugaugugaaa gcccccgcu caaccgggga gggucuuugg aaacugggag acuugagugc 660  
agaagaggag aguggaauuc caugugugagc gguuagaaugc guagauuuu ggaggaacac 720  
caguggcgaa ggcggcucuc uggucuguaa cugacgcuga ggcucgaaag cguggggagc 780  
aaacaggauu agauaccucg guaguccacg ccguaaacga ugagugcuaa guuguggagg 840  
guuuccgcc uucagugcug cagcuaacgc auuaagcacu ccgccugggg aguacgaccg 900  
caagguugaa acucaaaagga auugacgggg gcccgcaaa cgggugggagc augugguuu 960  
auucgaagca acgcgaagaa ccuuaaccag ucuugacauc cuuugaccac ucuagagaua 1020  
gagcuucccc uucgggggca aagugacagg uggugcaugg uugucgucag cucgugucgu 1080  
gagauguugg guuaagucc gcaacgagcg caaccuuuau uguuaguugc caucauuuag 1140  
uugggacuc uagcaagacu gccggugaca aaccggagga agguggggau gacgucaau 1200  
caucaugccc cuuaugaccu gggcuacaca cgugcuacaa ugggaaguac aacgagucgc 1260  
aaagucgca ggcuaagcua aucucuuaaa gcuucucuca guucggauug uaggcugcaa 1320  
cucgccuaca ugaagccgga aucgcuaqua aucgcggauc agcacgccgc ggugaauacg 1380  
uucccgggcc uuguacacac cggccgucac accacgagag uuuguaacac ccgaagucgg 1440  
ugagguaacc uuuuggagcc agccgccuaa gguugggauag augauugggg ugaagucgua 1500  
acaagguagc cguaucggaa ggugcggcug gauca 1535

<210> 24

<211> 1435

5 <212> DNA

<213> Tetragenococcus halophilus

<220>

<221> modified\_base

10 <222> (6)..(6)

<223> n e s a, c, g, o t

<220>

<221> modified\_base

15 <222> (53) .. (53)

<223> n e s a, c, g, o t

<220>

<221> modified\_base

20 <222> (253)..(253)

<223> n e s a, c, g, o t

ES 2 391 744 T3

<220>  
 <221> modified\_base  
 <222> (298)..(298)  
 <223> n es a, c, g, o t  
 5  
 <220>  
 <221> modified\_base  
 <222> (1094)..(1094)  
 <223> n es a, c, g, o t  
 10  
 <220>  
 <221> modified\_base  
 <222> (1182)..(1182)  
 <223> n es a, c, g, o t  
 15  
 <220>  
 <221> modified\_base  
 <222> (1426)..(1426)  
 <223> n es a, c, g, o t  
 20  
 <400> 24  
  
 aatacntgca agtcgaaacgc tgcttaagaa gaaacttcgg tttt ttctta agngggagtgg 60  
 cggacgggtg agtaaacacgt ggggaaccta tccatcagcg ggggrataaca cttggaaaca 120  
 ggtgctaata ccgcataatgg ctttttttca cctgaaagaa agct caaagg cgctttacag 180  
 cgtcactgat ggctggctcc gcggtgcatt agccagttgg tgaggtaacg gctcaccaaa 240  
 gcaacgatgc atngccgacc tgagaggggtg atcggccaca ctgggactga gacacggnc 300  
 agactcctac gggaggcagc agtagggaat ctteggcaat ggacgcaagt ctgaccgagc 360  
 aacgccgcgt gagtgaagaa ggttttcgga tcgtaaagct ctgt tgtcag caaagaacag 420  
 gagaaagagg aaatgctttt tctatgacgg tagctgacca gaaagccacg gctaactacg 480  
 tgccagcagc cgcggtaata cgtaggtggc aagcgttgtc cggatttatt gggcgtaaag 540  
 cgagcgcagg cggtgattta agtctgatgt gaaagccc c agctcaactg gggaggggtca 600  
 ttgaaactg gatcacttga gtgcagaaga ggagagtggga attccatgtg tagcggtgaa 660  
 atgcgtagat atatggagga acaccagtgg cgaaggcggc tctctggtct gtaactgacg 720  
 ctgaggctcg aaagcgtggg tagcaaacag gattagat ac cctggtagtc cacgccgtaa 780  
 acgatgagtg ctaagtgttg gagggttcc gcccttcagt gctgcagtta acgcattaag 840  
 cactccgcct ggggagtacg accgcaaggt tgaaactcaa aggaattgac gggggcccgc 900  
 acaagcggtg gagcatgttg ttaattcga agcaacgcga agaaccttac caggtcttga 960  
 catcctttga ccgccctaga gatagggttt ccccttcggg ggcaaagtga caggtggtgc 1020  
 atggttgtcg tcagctcgtg tcgtgagatg ttgggtta ag tcccgtaacg agcgcaacct 1080  
 ttattgttag ttgncagcat tgagttgggc actctagc aa gactgccggt gacaaaccgg 1140  
 aggaaggcgg ggatgacgtc aatcatcat gccctta tg anctgggcta cacacgtgct 1200  
 acaatgggaa gtacaacgag caagccaagc cgcaaggc ct agcgaatctc tgaaagcttc 1260  
 tctcagttcg gattgcaggc tgcaactcgc ctgcatga ag ccggaatcgc tagtaatcgc 1320  
 ggatcagcat gccgcgggtg atccgtccc gggccttgrta cacaccgccc gtcacaccac 1380  
 25 gagagtttgt aacacccaaa gtcgggtcgg caaccctt cg gggagncagc cgcct 1435  
  
 <210> 25

ES 2 391 744 T3

<211> 1452  
 <212> ARN  
 <213> Escherichia coli

5 <220>  
 <221> modified\_base  
 <222> (491)..(492)  
 <223> n es a, c, g, o u

10 <220>  
 <221> modified\_base  
 <222> (567)..(567)  
 <223> n es a, c, g, o u

15 <220>  
 <221> modified\_base  
 <222> (571) .. (571)  
 <223> n es a, c, g, o u

20 <220>  
 <221> modified\_base  
 <222> (852)..(853)  
 <223> n es a, c, g, o u

25 <400> 25

```

aguuuugauc a uggcucagau ugaacgcugg cggcagggccu aacacaugca agucgaacgg 60
uaacagggaac gagcuugcug cuuugcugac gaguggcgga cgggugagua augucuggga 120
aacugccuga uggaggggga uaaacuacugg aaacgguagc uaaauaccgca uaacgucgca 180
agaccaaaga gggggaccuu cgggccucuu gccaucggau guggccagau gggauuagcu 240
aguagguggg guaaaggcuc accuaggcga cgaucccuag cuggucugag aggaugacca 300
gccacacugg aacugagaca cgguccagac uccuacggga ggcagcagug gggaaauuug 360
cacaaugggc gcaagccuga ugcagccaug ccgcguguau gaagaaggcc uucggguugu 420
aaaguacuuu cagcggggag gaagggagua aaguuaauac cuuugcucau ugacguuacc 480
cgcagaagaa nnaccggcua acuccgugcc agcagccgcg guaaucgga gggugcaagc 540
guuaaucgga auuacugggc guaaagngca ngcaggcggg uuguuaaguc agaugugaaa 600
ucccggggcu caaccuggga acugcaucug auacuggcaa gcuugagucu cguagagggg 660
gguagaauuc caggguagc ggugaaugc guagagaucu ggaggaauc cgguggcgaa 720
ggcggccccc uggacgaaga cugacgcuca ggugcgaaag cgggggagc aaacaggauu 780
agauaccug guaguccag ccguaaacga ugucgacuug gagguugugc ccuugaggcg 840
uggcuuccgg annuaacgcg uuaagucgac cgcuggggga guacggccgc aagguaaaa 900
cucaaaugaa uugacggggg ccgcacaagc gguggagcau gugguuuau ucgaugcaac 960
gcgaagaacc uuaccugguc uugacaucca cggaauguuu cagagaugag aaugugccuu 1020
cgggaaccgu gagacaggug cugcauggcu gucgucagcu cguguuguga aauguugggu 1080
uaagucccg aacgagcgca acccuuaucc uuuguugcca ggguccggc cgggaacuca 1140
aaggagacug ccagugauaa acuggggaa gguggggaug acgucaguc aucauggccc 1200
uuacgaccag ggcuacacac gugcuacaau ggcgcuaaca aagagaagcg accucgag 1260
agcaagcggg ccucauaaag ugcgucguag uccggauugg agucugcaac ucgacuccau 1320
gaagucggaa ucgcuagua ucguggauca gaaugccaag gugaauacgu ucccgggccu 1380
uguacacacc gcccgucaca ccaugggagu ggguuugcaaa agaaguaggu agcuuaaccu 1440
ucgggagggc gc 1452
    
```



ES 2 391 744 T3

<210> 26  
 <211> 1541  
 <212> ARN  
 <213> Salmonella bovis

5

<400> 26

```

aaauugaaga guuugaucuau ggcucagauu gaacgcuggc ggcaggccua acacaugcaa      60
gucgaacggg aacaggaaga agcuugcucg cugcugacga guggcggacg ggugcguaau     120
gucugggaaa cugccugaug gagggggaa acuaucggaa acgguggcua aucccgcaua     180
acgucgcaag accaaagagg gggaccucca ggocucucc caucggaugu gcccagaugg     240
gauuagcuag uuggugaggu aacggcucac caagggcagc aucccuagcu ggucugagag     300
gaugaccagc cacacuggaa cugagacacg guccagacuc cuacgggagg cagcaguggg     360
gaauauugca cagugugcgc aagccugaug cagccaugcc gcucguauga agaaggccuu     420
cggguuguaa aguacuuuca ggggggagga aggugugug guuaauaacu gcagcaauug     480
acguuaccgg cagaagaagc accggcuaac uccgugccag caggccgggu aaucggagg     540
gugcaagcgu uaaucggaa uacugggcg uaaagcgcag caggcgguuu guuaagucag     600
augugaauc cccgggcuca accugggaac ugcaucugau acuggcaagc uugagucucg     660
uagagggggg uagaauucca ggugugcgg ugaauugcgu agagaucugg aggaauaccg     720
guggcgaagg cggccccug gacgaagacu gacgcucagg ugcgaaagcg uggggagcaa     780
acaggauuag auaccuggu aguccacgcc guaaacgaug ucuacuugga gguugugccc     840
uugagggcug gcuuccggag cuaacgcguu aaguagaccg ccuggggagu acggccgcaa     900
gguaaaaacu caaaugaauu gacgggggccc gcacaagcg guggagcaug ugguuuaau     960
ccaugcaacu cuaagaaccu uaccugguca ugcauccac a gaacuuucc agagaugaga    1020
cugugccuuc gggaacugug agacaggugc ugcauggcug ucgucagcuc guguugugaa    1080
auguuggguu aagucccgca acgagcgcaa ccuuauccu uuguugccag cgguccggcc    1140
gggaacucua aggagacugc cagugauaaa cuggaggaag guggggauga cgucaaguca    1200
ucaugcccuu uacgaccagg gcuaacacag ucguacaau ggcgcauaca agagaagcga    1260
ccucgcgaga gcaagcggac cucauaaagu gcgucguagu ccggauugga gucugcaacu    1320
cgacuccaug aagucggaa cgcuaaguau cguggaucag aaugccacgg ugaauacguu    1380
cccgggccuu guacacaccg cccgucacac caugggagug gguugcaaaa gaaguaggua    1440
gcuuaaccuu cgggagggcg cuuaccacuu ugugauucau gacuggggug aagucguaac    1500
aagguaacgg uaggggaacc ugcgguugga ucaccuccu a                                1541
    
```

10

<210> 27  
 <211> 1488  
 <212> ARN  
 <213> Shigella boydii

15

<400> 27

ES 2 391 744 T3

uggcucagau ugaacgcugg cggcaggccu aacacaugca agucgaacgg uaacaggaag 60  
 cagcuugcug uuucgcugac gaguggcgga cgggugagua augucuggga aacugccuga 120  
 uggaggggga uaacucugg aaacgguagc uaauaccgca uaacgucgca agaccaaaga 180  
 gggggaccuu cgggcccucu gccaucggau gugcccagau gggauuagcu uguuggugg 240  
 guaacggcuc accaaggcga cgaucccuaug cuggucugag aggaugacca gccacacugg 300  
 aacugagaca cgguccagac uccuacggga ggcagcagug gggauuuuug cacaaugggc 360  
 gcaagccuga ugcagccaug ccgcguguau gaagaaggcc uucggguugu aaaguacuuu 420  
 cagcggggag gaagggagua aaguuuuuac cuuugcucu ugacguuau cgcagaagaa 480  
 gcaccggcua acuccgugcc agcagccgcg guauuacgga gggugcaagc guuuuacgga 540  
 auuacugggc guaaaagcga cgcaggcggg uuuuuaguc agauguaaa uccccgggcu 600  
 caaccuggga acugcaucug auacuggcaa gcuugagucu cguagagggg gguagaauuc 660  
 cagguguagc ggugaaaugc guagagau cu ggaggauuac cgguggcgaa ggcggcccc 720  
 uggacgaaga cugagcucu gggugcgaag cggggggagc aaacaggauu agauaccucg 780  
 guaguccacg ccguaaaagc ugcgacuug gagguugugc ccuugagggc uggcuuccgg 840  
 agcuaacgcg uuaagucgac cggcggggga guacggccgc aagguuuuuu cucaaavga 900  
 uugacggggg cccgcacaag cgguggagca ugugguuuu uucgaugcaa cgcgaagaac 960  
 cuuaccuggu cuugacaucc acggaaguuu ucagagauga gaauugccu ucgggaaccg 1020  
 ugagacaggu gcugcauggc ugucgucagc ucguugugug aaauugggg uuaagucccg 1080  
 caacgagcgc aaccuuuau cuuuuuugc agcgguccgg cggggaacuc aaaggagacu 1140  
 gccagugaua aacuggagga agggggggau gacgucaagu caucauggcc cuuacgacca 1200  
 gggcuacaca cgugcuacaa uggcgcauac aaagagaagc gaccucgca gacgaagcgg 1260  
 accucauaaa gugcgucgua guccggauug gagucugcaa cucgacucca ugaagucgga 1320  
 aucgcuagua aucguggauc agaauccac ggugaauacg uucccgggcc uuguacacac 1380  
 cggccgucac accaugggag uggguugcaa aagaaguag uagcuuacc uucgggaggg 1440  
 cguuaccac uuugugauuc augacugggg ugaagucgua acaaggua 1488

<210> 28  
 <211> 1471  
 5 <212> ARN  
 <213> Shigella dysenteriae

<220>  
 <221> modified\_base  
 10 <222> (1)..(2)  
 <223> n es a, c, g, o u

<220>  
 <221> modified\_base  
 15 <222> (1470)..(1470)  
 <223> n es a, c, g, o u

<400> 28

ES 2 391 744 T3

nnauugaaga guuugaucuau ggcucagauu gaacgcuggc ggcaggccua acacaugcaa 60  
 gucgaacggu aacaggaagc agcuugcugc uuugcugacg aguggcggac gggugaguaa 120  
 ugucugggaa acugccugau gggagggggau aacuacugga aacggguagcu aaauaccgcau 180  
 aacgucgcaa gaccaaagag ggggaccuuc gggccucuuug ccuacggauug ugcccagaug 240  
 ggaauagcua guagggggg uaauggcuca ccuagggcagc gaucuccuagc uggucugaga 300  
 ggaugaccag ccacacugga acugagacac gguccagacu ccuacgggag gggagcagugg 360  
 ggaauauugc acaaugggag caagccugau gcagccaugc gcguguaug aagaaggcuu 420  
 cggguuguaa aguacuuuca gcggggagga agggaguaaa guuaauaccu uugcucauug 480  
 acguuacccg cagaagaagc accggcuaac uccgugccag cagccgcggu aaucggagg 540  
 gugcaagcgu uaaucggau uacugggcgu aaagcgcacg caggcgguu guuaagucag 600  
 augugaauc cccgggcuca accugggaac ugcaucugau acuggcaagc uugagucucg 660  
 uagagggggg uagaauucca gguguagcgg ugaaaugcgu agagaucugg aggaauaccg 720  
 guggcgaag cggccccug gacgaagacu gacgcucagg ugcgaaagcg uggggagcaa 780  
 acaggauuag auaccuggu aguccacgce guaaacgaug ucgacuugga gguugugccc 840  
 uugagggcgu gcuuccggag cuaacgcguu aagucgaccg ccuggggagu accggccgcaa 900  
 gguuaaaacu caaaugaauu gacgggggccc gcacaagcgg uggagcaugu gguuuauuc 960  
 gaugcaacgc gaagaaccuu accugguuuu gacauccacg gaaguuuua gagaugagaa 1020  
 ugugccuucg ggaaccguga gacaggugcu gcauggcugu cgucagcucg uguugugaaa 1080  
 uguuggguua agucccgcaa cgagcgaac ccuuauccuu uguugccagc gguccggccg 1140  
 ggaacucaaa ggagacugcc agugauaaac uggaggaagg uggggaugac gucaagucau 1200  
 cauggcccuu acgaccaggg cuacacacgu gcuacaauug cgcauacaaa gagaagcagc 1260  
 cucgcgagag caagcggacc ucauaaagug cgucguaguc cggauuggag ucugcaacuc 1320  
 gacuccauga agucggauc gcuaguaaauc guggaucaga augccacggu gaauacguuc 1380  
 ccgggccuug uacacaccgc ccgucacacc augggagugg guugcaaaag aaguagguag 1440  
 cuuaacuucg gggggcgc uaccacuun u 1471

5 <210> 29  
 <211> 1468  
 <212> ARN  
 <213> Shigella flexneri

10 <220>  
 <221> modified\_base  
 <222> (1) .. (2)  
 <223> n es a, c, g, o u

15 <220>  
 <221> modified\_base  
 <222> (372)..(372)  
 <223> n es a, c, g, o u

20 <400> 29

ES 2 391 744 T3

nnauugaaga guuugaucu ggcucagauu gaacgcuggc ggcaggccua acacaugcaa 60  
 gucgaacggg aacaggaagc agcuugcugu uucgcugacg aguggcggac gggugaguaa 120  
 ugucugggaa acugccugau ggagggggau aacuaucgga aacgguagcu aaauccgcau 180  
 aacgucgcaa gaccaaagag ggggaccuuc gggccucuuug ccuacgggag ugcccagaug 240  
 ggauuagcua guagggggg uaacggcuca ccuagggcag gaucccuagc uggucugaga 300  
 ggaugaccag ccacacugga acugagacac gguccagacu ccuacgggag gcagcagugg 360  
 ggaauauugc anaaugggcg caagccugau gcagccaugc cgcguguaug aagaaggccu 420  
 ucggguugua aaguacuuc agcggggagg aagggaguaa aguaauacc uuugcuauu 480  
 gacguuacc gcagaagaag caccggcuua cuccgugcca gcagccggcg uaauacggag 540  
 gguucaagcg uuaaucggaa uuacugggcg uaaagcgcac gcagggcggu uguuaaguca 600  
 gaugugaaau ccccgggcuc aaccugggaa cugcaucuga uacuggcaag cuugagucuc 660  
 guagaggggg guagaauucc agguugagcg gugaauugcg uagagaucug gaggaauacc 720  
 gguugcgagc gcggcccccu ggacgaagac uacgcucagc gugcgaaagc gugggggagca 780  
 aacaggauua gauaccucgg uaguccacgc uguaaacgau gucgacuugg agguugugcc 840  
 cuugaggugu ggcuucggga cguaacgcgu uaagucgacc gccuggggag uacggccgca 900  
 agguuaaaac ucaaaugaau ugacgggggc cgcacaagcg guggagcaug ugguuuaau 960  
 cgauzcaacg cgaagaaccu uaccuggucu ugacauccac ggaaguuuuc agagaugaga 1020  
 augugccuuc gggaaaccgug agacaggugc ugcauggcug ucgucagcuc guguugugaa 1080  
 auguuggguu aagucccgca acgagcgcaa cccuuauccu uuguugccag cgguccggcc 1140  
 gggaaacuaa aggagacugc cagugauaaa cuggaggagc guggggauga cgucaaguca 1200  
 ucaugggccu uacgaccagc gcuacacacg ugcuaaaug gcgcuaucua agagaagcga 1260  
 ccucgcgaga gcaagcggac cucacaaagu gcgucguagu ccggauugga gucugcaacu 1320  
 cgacuccaug aagucggaau cgcuaaauu cguggaucag aaugccacgg ugaauacguu 1380  
 cccggggccu guacacaccg cucgucacac caugggagug gguuguaaaa gaaguaggua 1440  
 gcuaaacuuc gggagggcgc uuaccacu 1468

5 <210> 30  
 <211> 1541  
 <212> ARN  
 <213> Streptococcus bovis

10 <400> 30

ES 2 391 744 T3

agaguugau ccuggcucag gacgaacgcu ggcggcgugc cuaauacaug caaguagaac 60  
 gcugaagacu uuagcuugcu aaaguuggaa gaguugcgaa cgggugagua acgcuaggu 120  
 aaccugccua cuagcggggg auaacuauug gaaacgauag cuaauaccgc auaacagcau 180  
 uuaacacaug uuagaugcuu gaaaggagca auugcuucac uaguaaugg accugcguug 240  
 uauuagcuag uuggugaggu aaaggcucac caaggcgacg auacauagcc gaccugagag 300  
 ggugaucggc cacacuggga cugagacacg gccagacuc cuacgggagg cagcaguagg 360  
 gaauucucgg caaugggggc aaccucgacc gagcaacgcc gcgugaguga agaagguuuu 420  
 cggaucguaa agcucugnuug uaagagaaga acguguguga gaguggaaag uucacacagu 480  
 gacgguaacu uaccagaaag ggacggcuua cuacugcca gcagccggg uauuacguag 540  
 gucccgagcg uuugccggau uuauugggcg uaaagcgagc gcagccgguu uauuagucu 600  
 gaaguuaaag gcaguggcuu aaccuuguu cgcuuuggaa acuguuagac uugagugcag 660  
 aaggggagag uggauuucca ugugnagcgg ugaaaugcgu agauauaugg aggaacaccg 720  
 guggcgaagc cggcucucug gucuguaacu gacgcugagg cucgaaagcg uggggagcaa 780  
 acaggauuag auaccucggg aguccacgcc guaaacgaug agugcuaggu guuaggcccu 840  
 uuccggggcu uagugccgca gcuaacgcau uaagcacucc gccuggggag uacgaccgca 900  
 agguugaaac ucaaaggaa ugcggggggc ccgcacaagc gguggagcau gugguuuau 960  
 ucgaagcaac gcgaagaacc uuaccagguc uugacauccc gaugcuauuc cuagagauag 1020  
 gaaguucuuu cggaaacaucg gugacaggug gugcauggu gucgucagcu cgugucguga 1080  
 gauguugggu uaagucccgc aacgagcgca accccuauug uuaguugcca ucauuaguu 1140  
 gggcacucua gcgagacugc cgguaauaaa ccggaggaag guggggauga cgucaauca 1200  
 ucaugcccu uaugaccggg gcuaacacag ugcuaaaug guugguacaa cgagucgca 1260  
 gucggugacg gcaagcaau ucuuuaagc caaucucagu ucggauugua ggcugcaacu 1320  
 cgccuacaug aagucggaau cgcuaaguau cgcggaucag cacgcccggg ugaauacguu 1380  
 cccgggccuu guacacaccg cccgucacac cagcagaguu uguaacacc cgaagucggug 1440  
 agguaacuu uuaggagcca gccgccuaag gugggauaga ugauuggggu gaagucguaa 1500  
 caagguagcc guaucggaag gucggcugg aucaccuccu u 1541

5 <210> 31  
 <211> 1492  
 <212> ARN  
 <213> Streptococcus infantarius

10 <400> 31

ES 2 391 744 T3

gcucaggacu aacgcuggcg gcgugccuaa uacaugcaag uagaacgcug aaaacuuuag 60  
 cuugcuaaaag uuugaagagu ugcgaacggg ugaguaacgc guagguaacc ugccuacuag 120  
 cggggggaaua cuauuggaaa cgauagcuaa uaccgcuaaa cagcauuuaa cccauguuag 180  
 augcuugaaa ggagcaauug cuucacuagu agauggaccu gcguguuuuu agcuaguugg 240  
 ugagguaacg gcucaccaag gcgacgauac auagccgacc ugagagggug aucggccaca 300  
 cugggacuga gacacggccc agacuccuac gggagggcagc aguagggaa cuucggcaau 360  
 gggggcaacc cugaccgagc aacgccgcgu gagugaagaa gguuuucgga ucguaaagcu 420  
 cuguuguaag agaagaauug gugugagagu ggaaaguuca cacagugacg guaacuuacc 480  
 agaaagggac ggcuaacuac gugccagcag ccgcgguuuu acguaggucc cgagcguugu 540  
 ccggauuuuu ugggfcuaaa gcgagcgcag gcgguuuuuu aagucugaag uuaaaggcag 600  
 uggcuuaacc auuguucgcu uuggaaacug uuagacuuga gugcagaagg ggagagugga 660  
 auuccaugug uagcggugaa augcguagau auaugggaga acaccggugg cgaaagcggc 720  
 ucucugggucu gnaacugacg cugagggcucg aaagcguggg gagcaaacag gauuagauac 780  
 ccugguaguc cacgccguaa acgaugagug cuagguguaa ggcccuucc ggggcuuagu 840  
 gccgcagcua acgcauuuag cacuccgcu ggggaguacg accgcaaggu ugaaacuaa 900  
 aggaauugac gggggcccg cacaagcggug gagcaugug uuuuuuucga agcaacgcga 960  
 agaaccuuac caggucuuga caucccgaug cuauuccuag agauagggaag uuucuuucgga 1020  
 acaucgguga cagguggugc augguugucg ucagcucgug ucgugagaug ugggguaag 1080  
 ucccgcaacg agcgcaacc cuauuguuag uugccaucu uaaguugggc acucuagcga 1140  
 gacugccggu aauaaaccgg aggaaggugg ggauagcugc aaaucaucau gcccuuuag 1200  
 accugggcu cacacgugcu acaaugguug guacmacgag ucgcgagucg gugacggcaa 1260  
 gcaaaucucu uaaagccaau cucaguucg auuguaggcu gcaacucgcc uacaugaagu 1320  
 cggaaucgcu aguaaucgag gauragcagc ccgcgugaa uacguucccg ggccuuuag 1380  
 acaccgccc ucacaccagc agauuuugua acaccggaag ucgugaggu aaccuuuag 1440  
 gagccagccg ccuaaggugg gauagaugau uggggugaag ucguuacaag gu 1492

5 <210> 32  
 <211> 1487  
 <212> ARN  
 <213> Streptococcus salivarius

10 <220>  
 <221> modified\_base  
 <222> (939)..(939)  
 <223> n es a, c, g, o u

15 <220>  
 <221> modified\_base  
 <222> (1059)..(1059)  
 <223> n es a, c, g, o u

20 <400> 32

ES 2 391 744 T3

uuuaaugaga guuugauccu ggCucaggac gaacgcuggc ggcgugccua auacaugcaa 60  
 guagaacgcu gaagagagga gcuugcucuu cuuggaugag uugcgaacgg gugaguaacg 120  
 cguagguaac cugccuugua gCGggggaua acuaauuggaa acgauagcua auaccgcaua 180  
 acaauggaug acccauguca uuuaauugaa aggggcaaa ugcuccacuac aagauggacc 240  
 ugcguugua uagcuaguag guGagguaac ggcucaccua ggcgacgava cauagccgac 300  
 cugagaggggu gaucggccac acUgggacug agacacggcc cagacuccua cgggagggcag 360  
 caguagggaa ucuucggcaa uGGggggcaac ccugaccgag caacgcccgc ugagugaaga 420  
 agguuuucgg aucguaaagc ucuuuuguaa gucaagaacg agugugagag uggaaaguuc 480  
 acacugugac gguagcuuac caGaaaggga cggcuaacua cgugccagca gcccgguua 540  
 uacguagguc ccgagcguug uCCggauuaa uugggcuua agcgagcgca ggcgguuga 600  
 uaagcugaa guuaaaggcu guGgcucaac cauaguucgc uuuggaaacu gucaaaauug 660  
 agugcagaag gggagagugg aaauccaugu guagcguga aaugcguaga uauaugGagg 720  
 aacaccggug gcgaaagcgg cucucugguc uguaacugac gcugaggcuc gaaagccGugg 780  
 ggagcgaaca ggauuagaa cccugguagu ccacgcccua aacgaugagu gcuagguuu 840  
 ggauccuuuc cgggauucag ugccgcagcu aacgcauuaa gcacuccgcc uggggagGuac 900  
 gaccgcaagg uugaaacuca aaggaauuga cgggggcccng cacaagcggg ggagcauug 960  
 guuaaaucg aagcaacgcg aagaaccua ccaggucuuu acauccgau gcuauuucua 1020  
 gagauagaa guuaucucgg uacaucggug acaggugng caugguuguc gucagcucgu 1080  
 gucgugagau guuggguuaa gucccgaac gagcgcaacc ccuauguaa guugccauca 1140  
 uucaguuggg cacucugcg agacugccgg uaauaaaccg gaggaaggug gggauagcgu 1200  
 caaaucua ugccccuuu gaccugggcu acacacgugc uacaugguu gguacaacga 1260  
 guugcgaguc ggugacggca agcuauuc uuaaagccaa ucucaguucg gauuguaaggc 1320  
 ugcaacucgc cuacaugaag ucggaaucgc uaguaaucgc ggauagcac gcccggguga 1380  
 auacguucc gggccuugua cacaccgcc gucacaccac gagaguugu aacaccGaa 1440  
 gucgugagg uaaccuuug gagccagccg ccuaaggugg gauagau 1487

5 <210> 33  
 <211> 1540  
 <212> ARN  
 <213> Streptococcus thermophilus

10 <220>  
 <221> modified\_base  
 <222> (130)..(130)  
 <223> n es a, c, g, o u

15 <400> 33

ES 2 391 744 T3

agaguugau ccuggcucag gacgaacggu ggccggcgugc cuaauacaug caagua gaac 60  
 gcugaagaga ggagcuugcu cuucuuuggau gaguugcgaa cgggugagua acgcgu aggu 120  
 aaccugccun guagcggggg auaacuauug gaaacgauag cuaauaccgc auaaca augg 180  
 augacacaug ucauuuuuuu gaaaggggca auugcuccac uacaagaugg accugc guug 240  
 uauuagcuag uaggugaggu aauggcuuac cuaggcgacg auacauagcc gaccug agag 300  
 ggugaucggc cacacuggga cugagacacg gccagacuc cuacgggagg cagcagruagg 360  
 gaaucuucgg caaugggggc aaccucgacc gagcaacgcc gcgugacuga agaaggruuu 420  
 cggaucguaa agcucuguuu uaagucaaga acggguguga gaguggaaag uucaca cagu 480  
 gacgguaágc uaccagaaag ggacggcuua cuacgugcca gcagccgagg uaauacguag 540  
 gucccggagc uuguccggau uuauugggcg uaaagcgagc gcaggcggu ugaauagcu 600  
 gaaguuaag gcuguggcuc aaccuauugu cgcuuuggaa acugucuaac uugagugcag 660  
 aaggggagag uggaauucca uguguagcgg ugaaauccgu agauauaug aggaacaccg 720  
 guggcgaaa gggcucucug gucuguaacu gacgcugagg cucgaaagcg uggggagcga 780  
 acaggauuag auaccucggg aguccacgcc guaaacgaug agugcuaggu guuggaucu 840  
 uuccgggauu cagugccgca gcuaacgcau uaagcacucc gccuggggag uacgaccgga 900  
 agguugaaac ucaaaggaau ugacggggcc cgcacaagcg guggagcaug ugguuuaau 960  
 cgaagcaacg cgaagaaccu uaccaccucu ugacaucccg augcuuuuc uagagauaga 1020  
 aaguuacuuu gguacaucgg ugacaggugg ugcaugguug ucgucagcuc gugucgugag 1080  
 auguuggguu aagucccgca acgagcgcaa ccccuauugu uaguugccau cauucaguug 1140  
 ggcaucucua gtagacugcc gguaauaac cggagggaagg uggggaugac gucaaaucua 1200  
 caugcccuu augaccuggg cuacacacgu gcuaacaugg uugguacaac gaguugcgag 1260  
 uccgugacgg cgagcuauuc ucuuuaagcc aaucucaguu cggauuguag gcugcaacuc 1320  
 gccuacauga agucggaauc gcuaquaau gcggaucagc acgcccggu gaauacguuc 1380  
 cggggccuug uacacaccgc ccgucacacc acgagaguuu gaaacacccg aagucgguga 1440  
 gguaaccuuu uggagccagc cgccuaaggu gggacagaug auugggguga agucguaaca 1500  
 agguagccga uccggaaggu gcggcuggau caccuccuuu 1540

5 <210> 34  
 <211> 1332  
 <212> ARN  
 <213> Fusobacterium alocis

10 <220>  
 <221> modified\_base  
 <222> (76)..(76)  
 <223> n es a, c, g, o u

15 <220>  
 <221> modified\_base  
 <222> (80) .. (80)  
 <223> n es a, c, g, o u

20 <220>  
 <221> modified\_base  
 <222> (258) .. (258)  
 <223> n es a, c, g, o u



ES 2 391 744 T3

<220>  
 <221> modified\_base  
 <222> (261)..(261)  
 <223> n es a, c, g, o u  
 5  
 <220>  
 <221> modified\_base  
 <222> (374) .. (374)  
 <223> n es a, c, g, o u  
 10  
 <220>  
 <221> modified\_base  
 <222> (812)..(812)  
 <223> n es a, c, g, o u  
 15  
 <220>  
 <221> modified\_base  
 <222> (847)..(847)  
 <223> n es a, c, g, o u  
 20  
 <220>  
 <221> modified\_base  
 <222> (999)..(999)  
 <223> n es a, c, g, o u  
 25  
 <220>  
 <221> modified\_base  
 <222> (1076)..(1076)  
 <223> n es a, c, g, o u  
 30  
 <220>  
 <221> modified\_base  
 <222> (1174)..(1174)  
 <223> n es a, c, g, o u  
 35  
 <400> 34

```

ggguggcgga cgggugagua acgcguaaag aacuugccuc acagauaggg acaacauua 60
gaaaugaaug cuaaunccun auuuuugaa aauagggcau ccuaauuaa ugaaagcuaa 120
augcgcugug agagagcuuu gcgucccauu agcuaguug agagguaacg gcucaccaag 180
gcgauugaug guagccggcc ugagagggug aacggccaca aggggacuga gacacggccc 240
uuacuccuac gggaggcngc nguggggaau auuggacaau ggaccgagag ucugauccag 300
caauucugug ugcacgauga aguuuuucgg aauguaaagu gcuuucaguu gggagaaaa 360
gaaugacggu accnacagaa gaagugacgg cuaaaucgu gccagcagcc gcgguaauc 420
gaaugucacg agcguaaucc ggauuuauug ggcguaaagc gcgcuaggu gguuauguaa 480
gucugaugug aaaaugcagg gcucaacucu guauugcguu ggaaacugug uaacuagagu 540
acuggagagg uaagcggaac uacaagugua gaggugaaau ucguagauau uuguaggaa 600
gccgaugggg aagccagcuu acuggacaga uacugacgcu aaagcgcgaa agcgugggua 660
  
```

ES 2 391 744 T3

gcaaacagga uuagauaccc ugguaugucca cgccguaaac gaugauuacu agguuguuggg 720  
 ggucgaaccu cagcgcccaa gcuaacgcga uaaguaaucc gccuggggag uacguacgca 780  
 aguaugaaac ucaaaggaau ugacggggac cngcacaagc gguggagcau gugguuuau 840  
 ucgacgnaac gcgaggaacc uuaccagcgu uugacauuu aggaaugaga uacgagauua 900  
 ucagugucuu cggggaaacc uaaagacagg uggugcaugg cugucgucag cucgugucgu 960  
 gagauguugg guuaaguccc gcaacgagcg caaccuncu cguauguuac caucauuuag 1020  
 uuggggacuc augcgauacu gccugcgaug agcaggagga agggggggaug gacgnaagu 1080  
 caucaugccc cuuauacgcu gggcuacaca cgugcuacaa uggguagaac agagaguugc 1140  
 aaagccguga gguggagcua aucucagaaa acunuucua guucggauug uacucugcaa 1200  
 cucgaguaca ugaaguugga aucgcuagua aucgcgauc agcaaugucg cggugaauac 1260  
 guucucgggu cuuguacaca ccgccguca caccacgaga guugguugca ccugaaguag 1320  
 caggccuaac cg 1332

<210> 35

<211> 1461

5 <212> ARN

<213> Fusobacterium nucleatum

<220>

<221> modified\_base

10 <222> (288)..(288)

<223> n es a, c, g, o u

<220>

<221> modified\_base

15 <222> (637)..(637)

<223> n es a, c, g, o u

<220>

<221> modified\_base

20 <222> (744)..(746)

<223> n es a, c, g, o u

<220>

<221> modified\_base

25 <222> (941)..(941)

<223> n es a, c, g, o u

<400> 35

auugaacgaa gaguuugauc cuggcucagg augaacgcug acagaaugcu uaacacaugc 60  
 aagucuaacu gaauuugggu uuuuuuacu cgauuugggu ggcggacggg ugaguaacgc 120  
 guaaagaacu ugccucacag cuagggacaa cauuuggaaa cgaaugcuaa uaccuaauu 180

ES 2 391 744 T3

uaugauuuaa gggcauccua gaauuaugaa agcuauaugc gcugugagag agcuuugcgu 240  
 cccauuagcu aguuaggagag guaacggcuc accaaggcaa ugaugggnag ccggccugag 300  
 agggugaacg gccacaaggg gacugagaca cggcccuac uccuacggga ggcagcagug 360  
 gggaaauuug gacaauaggac cgagagucug auccagcaau ucugugugca cgaugacguu 420  
 uuucggaauug uaaagugcuu ucaguuggga agaaaaaaau gacgguacca acagaagaag 480  
 ugacggcuua auacgugcca gcagccgagg uaaauacguau gucacgagcg uuauccggau 540  
 uuauuggggcg uaaagcgcg cuaggugguu auguaagucu gaugugaaa ugcagggcuc 600  
 aacucugua ugcguuggaa acuguguaac uagaguncug gaaagguaag cggaaacuaca 660  
 aguguaagag ugaaauucgu agauuuugu aggaaugccg auggggaagc cagcuuacug 720  
 gacagauacu gacgcugaag cgcnnnagcg ugguagcaa acaggauuag auaccuggu 780  
 aguccacgcc guaaacgaug auuacuaggu guuggggguc gaaccucagc gcccaagcaa 840  
 acgcgauaag uauuccgccu ggggaguacg uacgcaagua ugaaacucua aggaaugac 900  
 ggggacccgc acaagcggug gagcaugugg uuuaauucga ngcaacgcga ggaaccuac 960  
 cagcguuuga caucuuaagga augagacaga gauguuucag ugucccuucg gggaaaccua 1020  
 aagacaggug gugcauggcu gucgucagcu cgugucguga gauguugggu uaaguccgc 1080  
 aacgagcgca acccuuucg uauguuacca ucauuuaguu ggggacucau gcgauacugc 1140  
 cuacgaugag uaggaggaag guggggaua cgucaaguca ucaugcccu uauacgcugg 1200  
 gcuacacacg ugcuaacaug gguagaacag agaguugcaa agccgugagg uggagcuau 1260  
 cucagaaaac uauucuuagu ucggauugua cucugcaacu cgaguacaug aaguuggaa 1320  
 cgcuaaguau cgcgaucag caaugucgag gugaauacgu ucucggguc uguacacacc 1380  
 gcccgucaca ccacgagagu ugguugcacc ugaaguagca ggccuaaccg uaaggaggga 1440  
 uguuccgagg gugugauuag c 1461

<210> 36  
 <211> 1452  
 5 <212> ARN  
 <213> Fusobacterium varium

<220>  
 10 <221> modified\_base  
 <222> (256)..(256)  
 <223> n es a, c, g, o u

<220>  
 15 <221> modified\_base  
 <222> (271)..(271)  
 <223> n es a, c, g, o u

<220>  
 20 <221> modified\_base  
 <222> (320)..(320)  
 <223> n es a, c, g, o u

<220>  
 25 <221> modified\_base  
 <222> (422)..(422)  
 <223> n es a, c, g, o u

<220>  
 <221> modified\_base

ES 2 391 744 T3

<222> (509)..(509)  
<223> n e s a, c, g, o u

<220>  
5 <221> modified\_base  
<222> (683)..(683)  
<223> n e s a, c, g, o u

<220>  
10 <221> modified\_base  
<222> (844)..(844)  
<223> n e s a, c, g, o u

<220>  
15 <221> modified\_base  
<222> (891)..(891)  
<223> n e s a, c, g, o u

<220>  
20 <221> modified\_base  
<222> (923)..(926)  
<223> n e s a, c, g, o u

<220>  
25 <221> modified\_base  
<222> (998)..(998)  
<223> n e s a, c, g, o u

<220>  
30 <221> modified\_base  
<222> (1157)..(1158)  
<223> n e s a, c, g, o u

<220>  
35 <221> modified\_base  
<222> (1447)..(1447)  
<223> n e s a, c, g, o u

<400> 36  
40 **aacgaagagu uugauccugg cucaggauga acgcugacag aaugcuuac acaugcaagu 60**

ES 2 391 744 T3

cuacuugauc uucgggugaa gguggcggac gggugaguaa cgcguaaaga acuugccuaa	1 20
cagacuggga caacauuugg aaacgaaugc uaauaccgga uauuugacu gagucgcaug	1 80
auuugguuu gaaagcuaua ugcgcuguga gagagcuuug cgucccauaa guuaguuggu	2 40
gagguaacgg cucacnaaga cgaugauggg nagccggccu gagaggguga acggccacaa	3 00
ggggacugag acacggcccn yacuccuacg ggaggcagca guggggaaua uuggacaau	3 60
gaccaaaagu cugauccagc aaucugugu gcacgaugaa guuuucgga auguaaagug	4 20
cnuucaguug ggaagaaguc agugacggua ccaacagaag aagcgacggc uaaaucgug	4 80
ccagcagccg cgguaauacg uaugucgcna gcguuauccg gauuuauugg gcguaaagcg	5 40
cgucugggc guuuaguaag ucugauguga aaaugcgggg cucaaccccg uauugcguug	6 00
gaaacugcua aacuagagua cuggagaggu aggcggaacu acaaguguag aggugaaau	6 60
cguagauuu uguaggauug ccnaugggga agccagccua cuggacagau acugacgcu	7 20
aagcgcgaaa gcguggguag caaacaggau uagauaccu gguaguccac gccguaaacg	7 80
augauuacua ggugugggg gucgaaccuc agcgcceaag cuaacgcgau aaguauccg	8 40
ccungggagu acguacgcaa guaugaaacu caaaggaaau gacggggacc ngcacaagcg	9 00
guggagcaug ugguuuuuu cgnnnnaacg cgaggaaccu uaccagcguu ugacauccca	9 60
agaaguuuac agagauuuu ucgugccucu ucggaggnac uuggugacag guggugcaug	10 20
gcugucguca gcucgugucg ugagauuuug gguuaagucc cgcaacgagc gcaacccuu	10 80
ucguaugua ccaucauuua guuggggacu caugcgagac ugccugcgau gagcaggagg	11 40
aaggugggga ugacgunnag ucaucaugcc ccuuauacgc ugggcuacac acgugcuaca	12 00
auggguagua cagagagcug caaacucgag aggguaagcu aaucucauaa aacuaucuu	12 60
aguucggauu guacucugca acucgaguac augaaguugg aaucgcuagu aaucgcaau	13 20
cagcuauuu gcgguagaaua cguucucggg ucuuguacac accgcccguc acaccacgag	13 80
aguugguugc accugaagua acaggccuaa ccguaaggag ggauguuccg agggugugau	14 40
uagcganugg gg	14 52

<210> 37  
 <211> 20  
 5 <212> ADN  
 <213> Secuencia artificial

<220>  
 <223> Descripción de la secuencia artificial: Oligonucleótido sintético  
 10 <400> 37

agagttgat cctggctcag

20

<210> 38  
 <211> 18  
 <212> ADN  
 <213> Secuencia artificial

<220>  
 <223> Descripción de la secuencia artificial: Oligonucleótido sintético  
 <400> 38

gctgcctccc gtaggagt

## REIVINDICACIONES

1. Procedimiento para diagnosticar una enfermedad inflamatoria intestinal en un paciente, que comprende determinar el aumento o la disminución en comparación con una muestra de mucosa normal de *Moraxella sp.* del grupo *Pseudomonas*, *Comamonas sp.* del grupo *Acidovorax*, *Cryseobacterium sp.* del grupo I Cytophaga, *Enterococcus sp.* del grupo *Enterococcus*, *Bacteroides sp.* del grupo *Bacteroides*, *Propionibacterium sp.* del grupo *Propionibacterium*, *Ruminococcus sp.* del grupo de *Clostridium Coccoides* y *Bacteroides sp.* en una muestra de tejido de la mucosa del colon de dicho paciente, en el que
- 5
- 10 (a) la enfermedad inflamatoria intestinal es la enfermedad de Crohn y en el que un aumento de *Moraxella sp.*, *Comamonas sp.* y *Chryseobacter sp.* y una disminución de *Bacteroides sp.*, *Propionibacterium sp.*, *Enterococcus sp.* y *Ruminococcus sp.* indica la enfermedad de Crohn en dicho paciente, o en el que
- 15 (b) la enfermedad inflamatoria intestinal es la colitis ulcerosa y en el que un aumento de *Moraxella sp.*, *Comamonas sp.* y *Enterococcus sp.* y una disminución de *Bacteroides sp.*, *Propionibacterium sp.*, *Chryseobacter sp.* y *Ruminococcus sp.* indica colitis ulcerosa en dicho paciente.
2. Procedimiento según la reivindicación 1, que comprende además determinar el aumento o la disminución de una o más bacterias seleccionadas de entre el grupo constituido por *Bacteroides sp.*, *Ruminococcus sp.*, *Moraxella sp.* y *Comamonas sp.* en una muestra de lumen del paciente.
- 20
3. Procedimiento para diagnosticar pouchitis en un paciente con bolsa del íleon, que comprende detectar *Fusobacter sp.*, *Propionibacterium sp.* y por lo menos una de las bacterias del grupo de *E.coli* y/o *Shigella sp.* en una muestra de tejido de la mucosa del colon de dicho paciente, en el que un aumento de *Fusobacter sp.* y una
- 25 disminución de una o más de las bacterias del grupo de *E.coli* y/o *Shigella sp.* en comparación con una muestra de bolsa normal y la ausencia de *Propionibacterium sp.* es diagnóstico de pouchitis.
4. Procedimiento según cualquiera de las reivindicaciones 1 a 3, en el que las bacterias se detectan utilizando un anticuerpo específico para dichas bacterias.
- 30
5. Procedimiento según cualquiera de las reivindicaciones 1 a 3, en el que las bacterias son detectadas utilizando una sonda de oligonucleótidos específica para dichas bacterias.
6. Procedimiento según la reivindicación 5, en el que las bacterias son detectadas utilizando técnica de huella de ácidos nucleicos, RCP, secuenciación de nucleótidos, inmunotransferencia Southern, y/o micromatrices de ADN que comprenden una pluralidad de secuencias específicas para una o más de dichas bacterias.
- 35
7. Procedimiento según cualquiera de las reivindicaciones 5 o 6, en el que dicha sonda de oligonucleótidos hibrida con una pequeña subunidad de ARN ribosómico o con una gran unidad de ARN ribosómico.
- 40
8. Procedimiento según la reivindicación 7, en el que dicha sonda de oligonucleótidos hibrida con las regiones variables o las regiones de hélice de dicho ARN ribosómico.
9. Utilización de un procedimiento según cualquiera de las reivindicaciones 1 a 8, para pronosticar y/o para realizar el seguimiento del curso de una enfermedad inflamatoria intestinal en un sujeto.
- 45

HLA: Crohn frente a Referencias

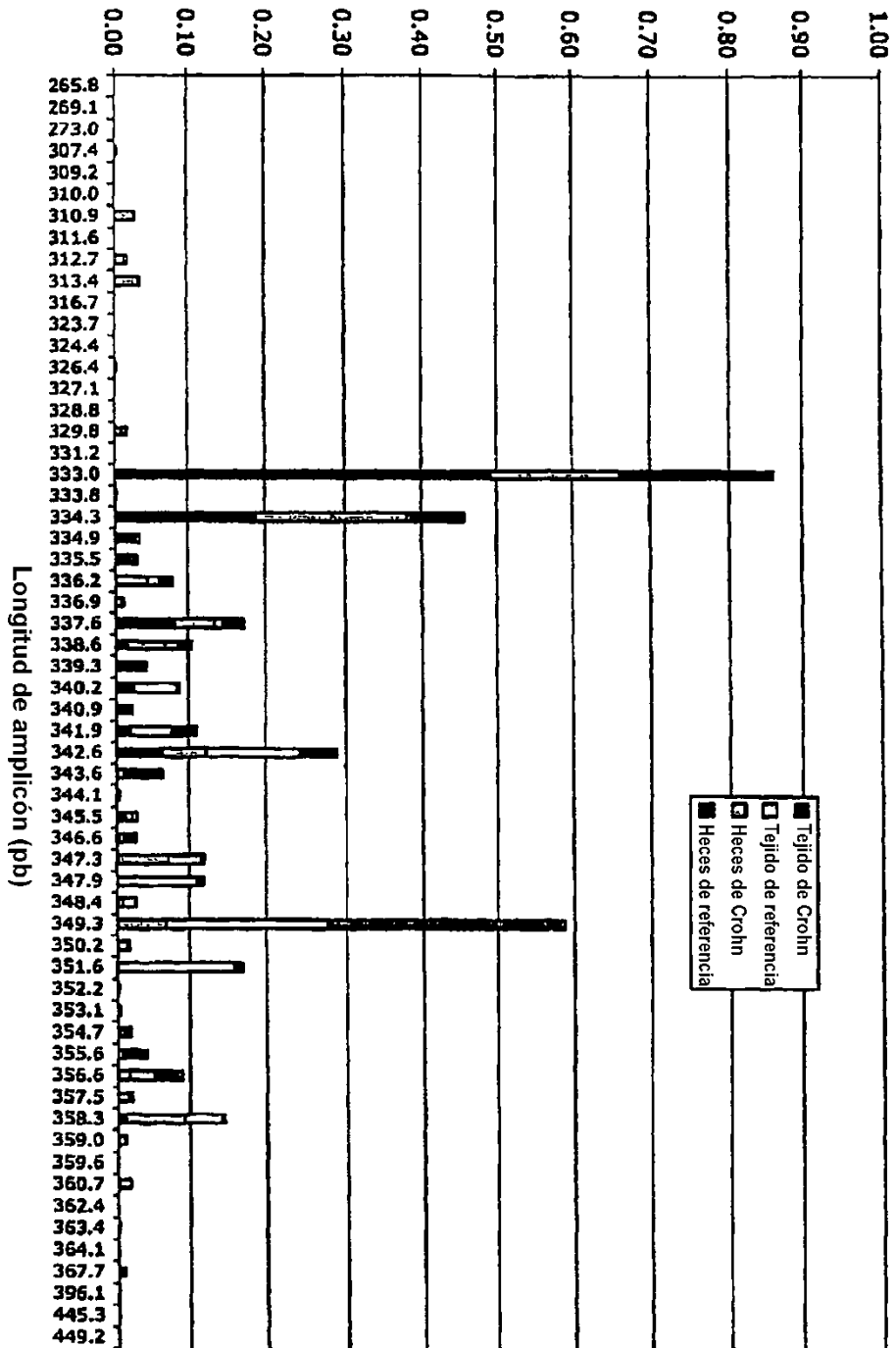


Figura 1: Comparación de tejido de Crohn y promedios de lumen

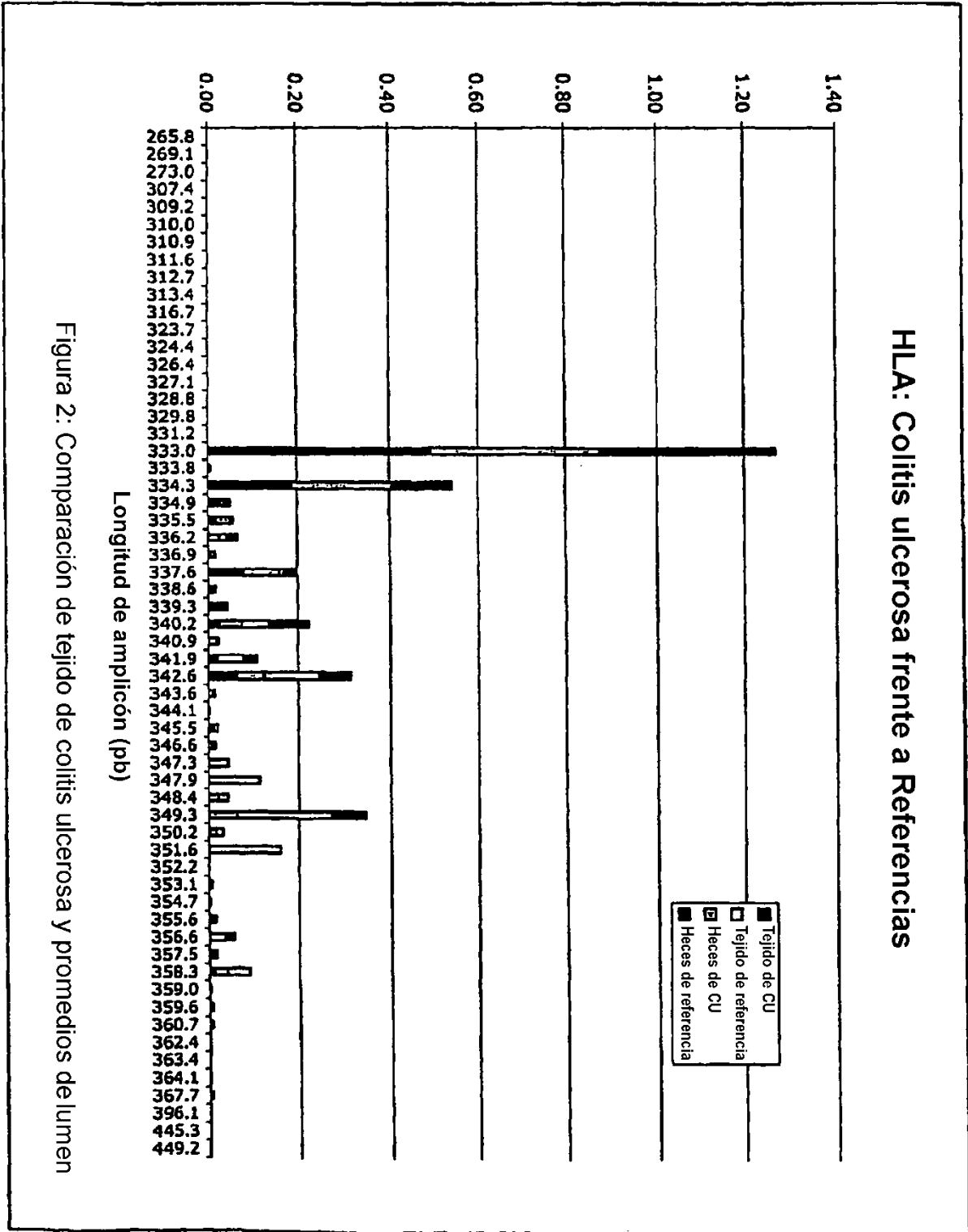


Figura 2: Comparación de tejido de colitis ulcerosa y promedios de lumen



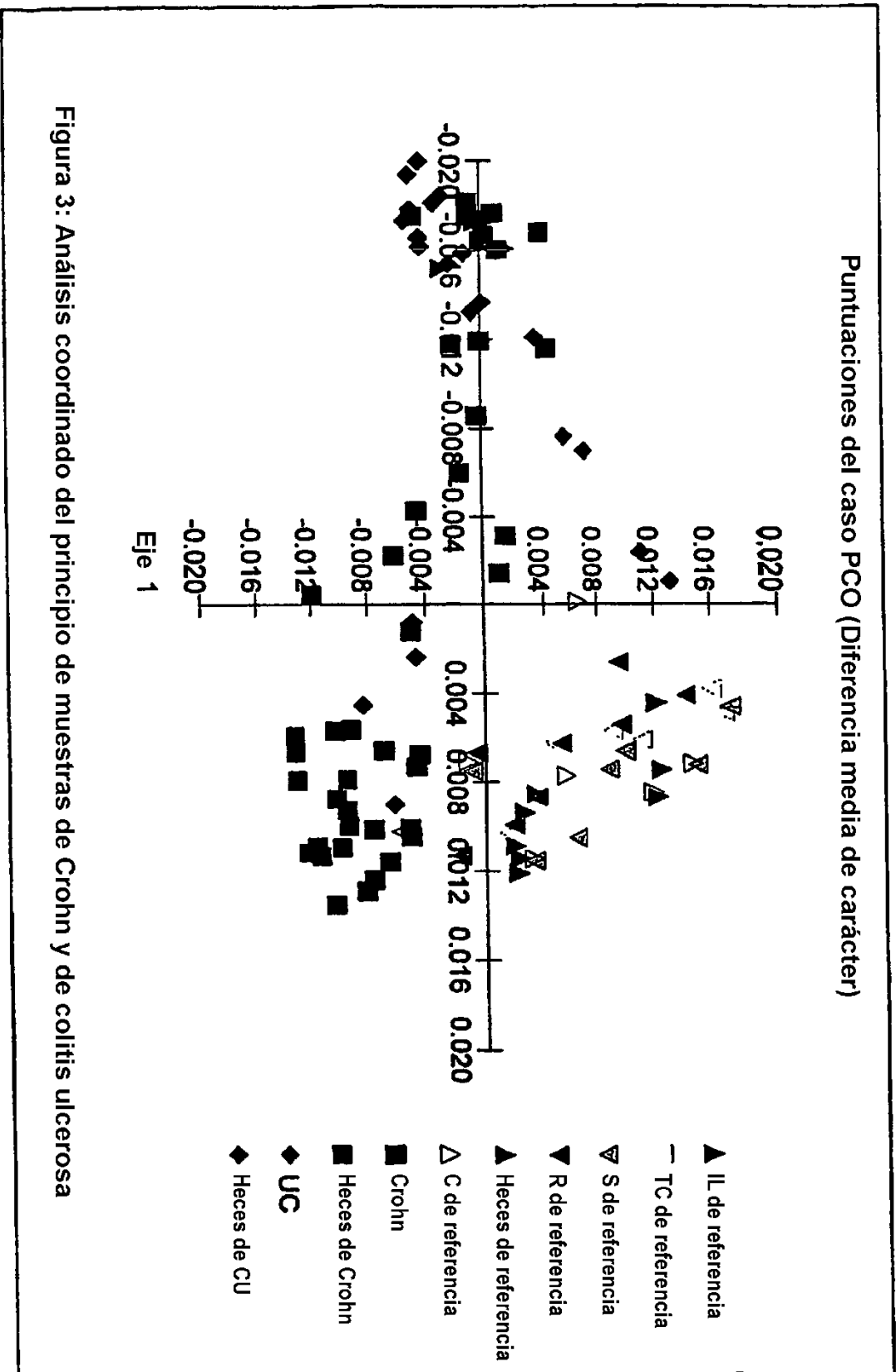
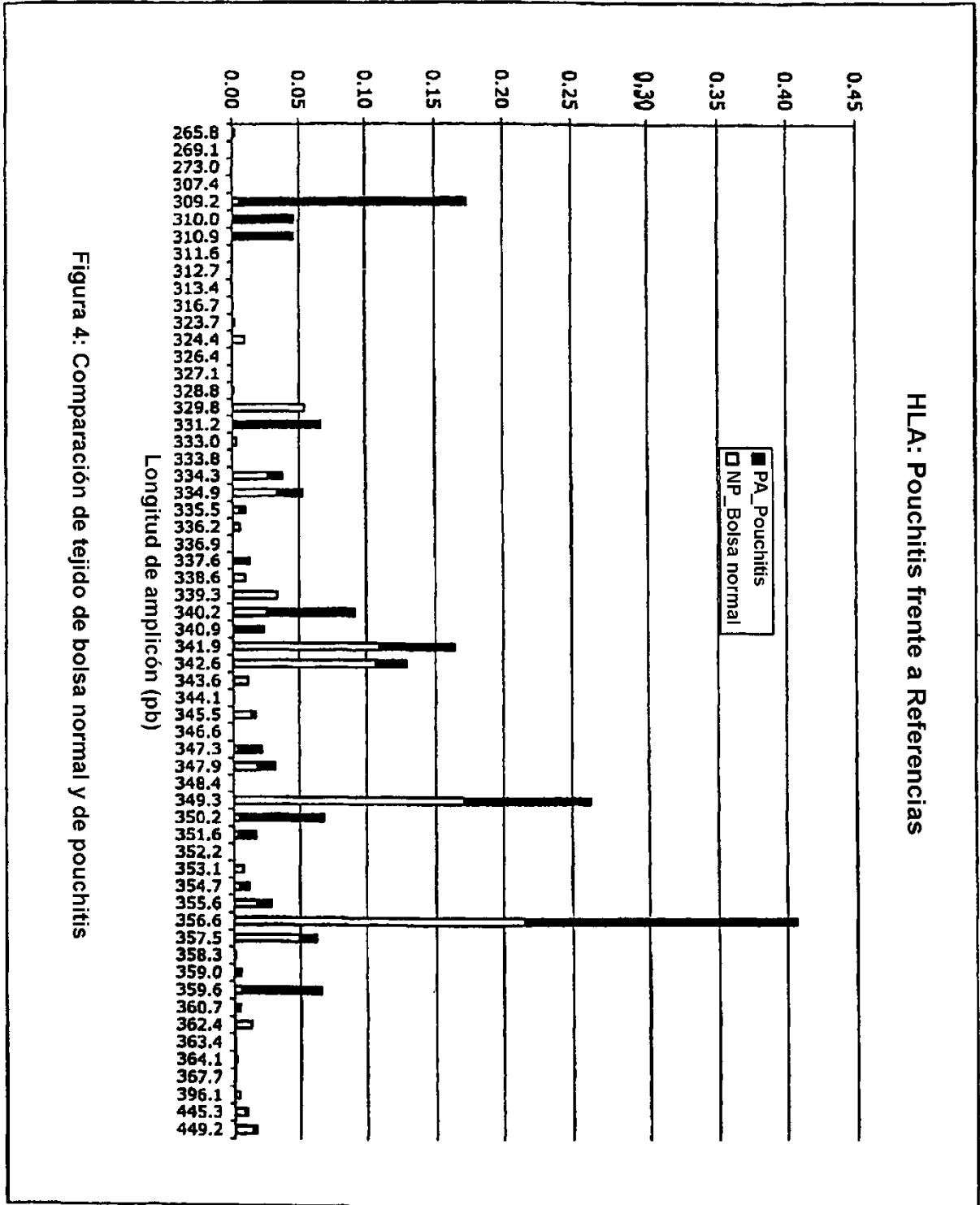


Figura 3: Análisis coordinado del principio de muestras de Crohn y de colitis ulcerosa



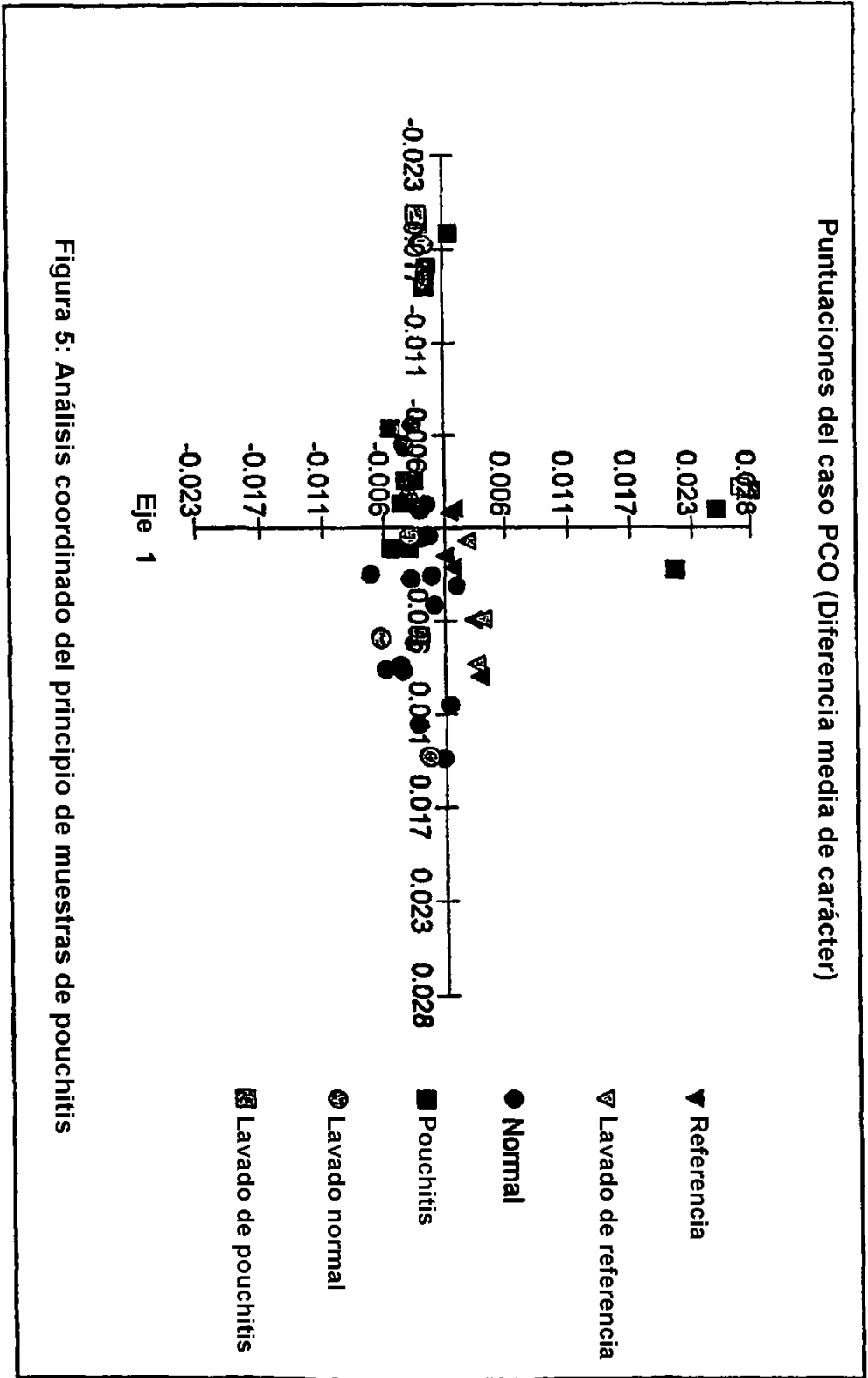


Figura 5: Análisis coordinado del principio de muestras de pouchitis

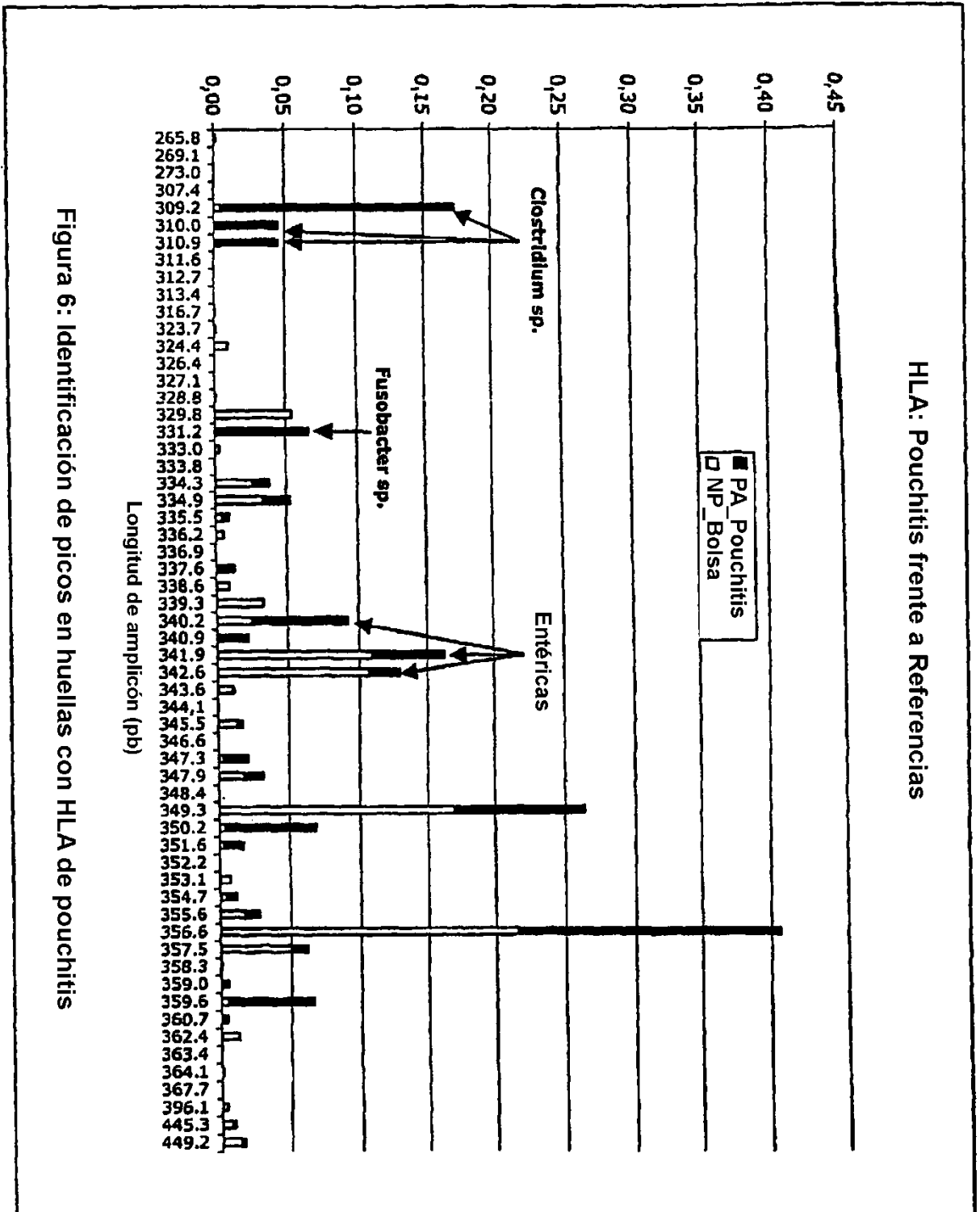


Figura 6: Identificación de picos en huellas con HLA de pouchitis