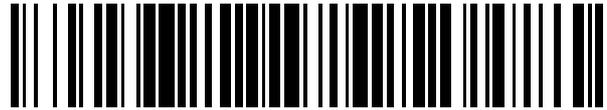


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 399 054**

51 Int. Cl.:

**C12Q 1/68**

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **05.12.2005 E 05813335 (6)**

97 Fecha y número de publicación de la concesión europea: **07.11.2012 EP 1828411**

54 Título: **Métodos para simplificar ácidos nucleicos microbianos mediante modificación química de citosinas**

30 Prioridad:

**03.12.2004 AU 2004906915**

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

**25.03.2013**

73 Titular/es:

**HUMAN GENETIC SIGNATURES PTY LTD  
(100.0%)**

**22 DELHI ROAD MACQUARIE PARK  
NSW 2113, AU**

72 Inventor/es:

**MILLAR, DOUGLAS, SPENCER y  
MIKLOS, GEORGE, GABOR, L.**

74 Agente/Representante:

**DE ELZABURU MÁRQUEZ, Alberto**

**ES 2 399 054 T3**

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

## DESCRIPCIÓN

Métodos para simplificar ácidos nucleicos microbianos mediante modificación química de citosinas

**Campo técnico**

5 La invención se refiere a análisis de detección de ácidos nucleicos para la detección de microorganismos. La invención también se refiere a métodos para el tratamiento químico de los ácidos nucleicos para reducir la complejidad de los genomas microbianos combinados con el uso de ligandos específicos para la detección microbiana.

**Técnica anterior**

10 En la actualidad se encuentran disponibles diversos procedimientos para la detección de moléculas de ácido nucleico específicas. Estos procedimientos dependen típicamente de la hibridación dependiente de la secuencia entre el ácido nucleico diana y sondas de ácido nucleico que pueden tener una longitud que varía de oligonucleótidos cortos (20 bases o menos) a secuencias de muchas kilobases (kb).

15 El método más ampliamente utilizado para la amplificación de secuencias específicas dentro de una población de secuencias de ácido nucleico es la reacción en cadena de la polimerasa (PCR) (Dieffenbach, C y Dveksler, G. eds. PCR Primer: A Laboratory Manual. Cold Spring Harbor Press, Plainview NY). En este método de amplificación, se utilizan oligonucleótidos, generalmente de 20 a 30 nucleótidos de longitud en cadenas de ADN complementarias y en cada extremo de la región que se va a amplificar, para cebar la síntesis de ADN sobre ADN de hebra sencilla desnaturalizado. Ciclos sucesivos de desnaturalización, hibridación de cebadores, y síntesis de hebras de ADN usando ADN polimerasas termoestables permiten la amplificación exponencial de las secuencias entre los  
20 cebadores. Las secuencias de ARN se pueden amplificar por medio de una primera copia utilizando la transcriptasa inversa para producir una copia de ADN complementario (ADNc). Se pueden detectar fragmentos de ADN amplificados mediante una variedad de métodos, incluyendo la electroforesis en gel, la hibridación con sondas marcadas, el uso de cebadores etiquetados que permiten la identificación posterior (por ejemplo, mediante un análisis con enzima ligada), y el uso de cebadores marcados con fluorescencia que dan lugar a una señal tras la  
25 hibridación con el ADN diana (por ejemplo, sistemas de Beacon y TaqMan).

Junto con la PCR, se ha desarrollado una variedad de técnicas distintas para la detección y amplificación de secuencias de nucleótidos específicas. Un ejemplo es la reacción en cadena de la ligasa (1991, Barany, F. et al., Proc. Natl. Acad. Sci. USA. 88, 189-193).

30 Otro ejemplo es la amplificación isotérmica que se describió por primera vez en 1992 (Walker GT, Little MC, Nadeau JG y Shank D. Isothermal in vitro amplification of DNA by a restriction enzyme/DNA polymerase system. PNAS 89: 392-396 (1992) y se denominó amplificación con desplazamiento de la hebra (SDA). Desde entonces, se han descrito otras diversas tecnologías de amplificación isotérmica incluyendo Amplificación Mediada por Transcripción (TMA) y Amplificación Basada en Secuencias de Ácido Nucleico (NASBA) que utilizan una ARN polimerasa para copiar secuencias de ARN pero no el ADN genómico correspondiente (Guatelli JC, Whitfield KM, Kwoh DY, Barringer KJ, Richmann DD y Gingeras TR. Isothermal, in vitro amplification of nucleic acids by a multienzyme reaction modeled after retroviral replication. PNAS 87: 1874-1878 (1990); Kievits T, van Gemen B, van Strijp D, Schukking R, Dircks M, Adriaanse H, Malek L, Sooknanan R, Lens P. NASBA isothermal enzymatic in vitro nucleic acid amplification optimized for the diagnosis of HIV-1 infection. J Virol Methods. 1991 Dec; 35(3):273-86).

40 Otras técnicas isotérmicas basadas en el ADN incluyen la amplificación mediante círculo rodador (RCA) en la que una ADN polimerasa prolonga un cebador dirigido a un molde circular (Fire A y Xu SQ. Rolling replication of short circles. PNAS 92: 4641-4645 (1995), Ramificación-Amplificación (RAM) que utiliza una sonda circular para la detección de la diana (Zhang W, Cohenford M, Lentricchia B, Isenberg HD, Simson E, Li H, Yi J, Zhang DY. Detection of *Chlamydia trachomatis* by isothermal ramification amplification method: a feasibility study. J Clin Microbiol. 2002 Jan; 40(1):128-32.) y más recientemente, amplificación de ADN isotérmica dependiente de helicasa (HDA), que utiliza una enzima helicasa para desenrollar las hebras de ADN en lugar de calor (Vincent M, Xu Y, Kong H. Helicase-dependent isothermal DNA amplification. EMBO Rep. 2004 Aug; 5(8):795-800.)

50 Recientemente, se han descrito métodos de amplificación isotérmica de ADN (Walker GT, Little MC, Nadeau JG y Shank D. Isothermal in vitro amplification of DNA by a restriction enzyme/DNA polymerase system. PNAS 89: 392-396 (1992). Las técnicas tradicionales de amplificación se basan en ciclos continuos de desnaturalización y renaturalización de las moléculas diana en cada ciclo de la reacción de amplificación. El tratamiento térmico del ADN da como resultado un cierto grado de cizallamiento de las moléculas de ADN, con lo que cuando el ADN es limitante como en el aislamiento de ADN a partir de un pequeño número de células de un blastocisto en desarrollo, o en particular en los casos en los que el ADN está ya en una forma fragmentada, por ejemplo en secciones de tejido, bloques de parafina y muestras de ADN antiguas, este ciclo de calentamiento-enfriamiento adicional podría dañar el  
55 ADN y dar como resultado la pérdida de señales de amplificación. Los métodos isotérmicos no se basan en la desnaturalización continua del ADN molde para producir moléculas de hebra sencilla que sirvan como moldes a

partir de una amplificación adicional, sino en la formación de mellas enzimáticas de moléculas de ADN mediante endonucleasas de restricción específicas a una temperatura constante.

La técnica denominada amplificación de desplazamiento de hebra (SDA) se basa en la capacidad de ciertas enzimas de restricción para formar una mella en la hebra de ADN hemi-modificada y la capacidad de una polimerasa carente de exonucleasa 5'-3' para ampliar y desplazar la hebra aguas abajo. La amplificación exponencial se consigue a continuación mediante el acoplamiento de reacciones efectoras y antisentido en las que el desplazamiento de la hebra de la reacción efectora sirve como molde para la reacción antisentido (Walker GT, Little MC, Nadeau JG y Shank D. Isothermal *in vitro* amplification of DNA by a restriction enzyme/DNA polymerase system. PNAS 89: 392-396 (1992). Dichas técnicas se han utilizado para la amplificación satisfactoria de *Mycobacterium tuberculosis* (Walker GT, Little MC, Nadeau JG y Shank D. Isothermal *in vitro* amplification of DNA by a restriction enzyme/DNA polymerase system. PNAS 89: 392-396 (1992), HIV-1, *Hepatitis C* and HPV-16 Nuovo G. J., 2000), *Chlamydia trachomatis* (Spears PA, Linn P, Woodard DL y Walker GT. Simultaneous Strand Displacement Amplification and Fluorescence Polarization Detection of *Chlamydia trachomatis*. Anal. Biochem. 247: 130-137 (1997).

El uso de SDA hasta la fecha ha dependido de nucleótidos modificados con fosforotioato con el fin de producir un dúplex de ADN-hemi-fosforotioato que en la hebra modificada sería resistente a la escisión por la enzima, dando como resultado la formación enzimática de mellas en lugar de la digestión para conducir la reacción de desplazamiento. Recientemente, sin embargo, se han diseñado varias enzimas "nickasa". Estas enzimas no cortan el ADN de la manera tradicional sino que producen una mella en una de las hebras de ADN. Las enzimas "nickasa" incluyen N.Alw1 (Xu Y, Lunnen KD y Kong H. Engineering a nicking endonuclease N.Alw1 by domain swapping. PNAS 98: 12990-12995 (2001), N.BstNB1 (Morgan RD, Calvet C, Demeter M, Agra R, Kong H. Characterization of the specific DNA nicking activity of restriction endonuclease N.BstNB1. Biol Chem. 2000 Nov; 381(11):1123-5.) y Mly1 (Besnier CE, Kong H. Converting Mly1 endonuclease into a nicking enzyme by changing its oligomerization state. EMBO Rep. 2001 Sep;2(9):782-6. Epub 2001 Aug 23). El uso de tales enzimas simplificaría de ese modo el procedimiento SDA.

Además, la SDA se ha mejorado mediante el uso de una combinación de una enzima de restricción estable al calor (Ava1) y una Exo-polimerasa estable al Calor (polimerasa Bst). Se ha demostrado que esta combinación aumenta la eficiencia de amplificación de la reacción de una amplificación de  $10^8$  veces a una amplificación de  $10^{10}$  veces de manera que es posible, usando esta técnica, la amplificación de moléculas únicas de copia única. El factor de amplificación resultante usando la combinación polimerasa/enzima estable al calor es del orden de  $10^9$  (Milia M. A., Spears P. A., Pearson R. E. y Walker G. T. Use of the Restriction Enzyme Ava1 and Exo-Bst Polymerase in Strand Displacement Amplification Biotechniques 1997 24:392-396).

Hasta la fecha, todas las técnicas de amplificación de ADN isotérmica requieren que la molécula ADN molde de doble hebra inicial se desnaturalice antes de la iniciación de la amplificación. Además, la amplificación sólo se inicia una vez de cada evento de cebado.

Para la detección directa, el ácido nucleico diana se separa muy comúnmente basándose en su tamaño mediante electroforesis en gel y se transfiere a un soporte sólido antes de la hibridación con una sonda complementaria a la secuencia diana (transferencia Southern y Northern). La sonda puede ser un ácido nucleico natural o análogo, tal como ácido peptidónucleico (PNA) o ácido nucleico cerrado (LNA) o ácido nucleico intercalante (INA). La sonda puede marcarse directamente (por ejemplo con  $^{32}\text{P}$ ) o se puede utilizar un procedimiento indirecto de detección. Los procedimientos indirectos por lo general dependen de la incorporación a la sonda de una "etiqueta" tal como biotina o digoxigenina y la sonda se detecta a continuación por medios tales como la conversión de sustrato unido a la enzima o la quimioluminiscencia.

Otro método para la detección directa de ácido nucleico que se ha utilizado ampliamente es hibridación "sándwich". En este método, una sonda de captura se acopla a un soporte sólido y el ácido nucleico diana, en disolución, se hibrida con la sonda unida. El ácido nucleico diana no unido se elimina mediante lavado y el ácido nucleico unido se detecta usando una segunda sonda que hibrida con las secuencias diana. La detección puede utilizar métodos directos o indirectos como se ha esbozado anteriormente. Los ejemplos de tales métodos incluyen el sistema de detección de señal "ADN ramificado", un ejemplo que utiliza el principio de hibridación sándwich (1991, Urdea, M. S., et al., Nucleic Acids Symp. Ser. 24, 197-200). Un área de rápido desarrollo que utiliza hibridación de ácido nucleico para la detección directa de secuencias de ácido nucleico es la de las micromatrices de ADN, (2002, Nature Genetics, 32, [Suplemento]; 2004, Cope, L.M., et al., Bioinformatics, 20, 323-331; 2004, Kendall, S.L., et al., Trends in Microbiology, 12, 537-544). En este procedimiento, las especies individuales de ácido nucleico, que pueden variar de oligonucleótidos cortos (típicamente, de 25 unidades en el sistema Affymetrix), a oligonucleótidos más largos, (típicamente de 60 unidades en Applied Biosystems y plataformas Agilent), a secuencias aún más largas, tales como clones de ADNc, se fijan a un soporte sólido en un patrón de cuadrícula o se sintetizan fotolitográficamente en un soporte sólido. A continuación se hibrida una población de ácido nucleico etiquetada o marcada con la matriz y se cuantifica el nivel de hibridación en cada punto en la matriz. Más comúnmente, se utilizan ácidos nucleicos marcados

radiactivamente o fluorescentemente (por ejemplo ARNc o ADNc) para la hibridación, aunque se pueden emplear otros sistemas de detección, tales como la quimioluminiscencia.

Un área de rápido crecimiento que utiliza hibridación de ácido nucleico para la detección directa de secuencias de ácido nucleico es la de las micromatrices de ADN (Young RA Biomedical discovery with DNA arrays. Cell 102: 9-15 (2000); Watson A New tools. A new breed of high tech detectives. Science 289:850-854 (2000)). En este procedimiento, las especies individuales de ácido nucleico, que puede variar de oligonucleótidos a secuencias más largas, tales como clones de ADN complementario (ADNc), se fijan a un soporte sólido en un patrón de cuadrícula. Una población de ácido nucleico etiquetado o marcado se hibrida a continuación con la matriz y se cuantifica el nivel de hibridación con cada punto de la matriz. Muy comúnmente, se utilizaron para la hibridación ácidos nucleicos marcados radiactivamente o fluorescentemente (por ejemplo, ADNc), aunque se emplearon otros sistemas de detección.

Los métodos tradicionales para la detección de microorganismos tales como bacterias, levaduras y hongos, incluyen el cultivo de microorganismos en medios de nutrientes selectivos clasificando a continuación los microorganismos basándose en el tamaño, la forma, la producción de esporas, caracteres tales como las reacciones bioquímicas o enzimática y las propiedades específicas de tinción (tales como la tinción de Gram) como se observa al microscopio óptico convencional. Las especies virales tienen que ser cultivadas en tejidos o células especializados, después clasificadas en función de su estructura y tamaño determinados mediante microscopía electrónica. Un inconveniente principal de tales técnicas es que no todos los microorganismos crecerán bajo condiciones convencionales de cultivo o de células que limitan la utilidad de tales enfoques. Con bacterias, por ejemplo, tales como *Neisseria meningitidis*, *Streptococcus pneumoniae* y *Haemophilus influenzae* (que causan todas meningitis y entre las cuales *N. meningitidis* causa tanto meningitis como meningococemia fulminante) las tres especies son difíciles de cultivar. Se examinan frascos de hemocultivo rutinariamente todos los días durante un máximo de siete días, y se requiere el subcultivo. *H. influenzae* requiere un medio especial que contiene dinucleótido de nicotinamida y adenina y hemina y el crecimiento en placas de agar chocolate. Los cultivos de sangre requieren caldo de tripticasa de soja o de infusión de cerebro-corazón y la adición de varios aditivos tales como polianetolsulfonato sódico. Para los microorganismos, tales como *Clostridium botulinum*, que provoca intoxicación alimentaria grave y síndrome de bebé hipotónico, la identificación de la toxina implica la inyección de extractos de alimentos o sobrenadantes de cultivo en ratones y la visualización de los resultados después de 2 días. Además, el cultivo del microorganismo potencial en medios especiales lleva una semana. La enterotoxina de *Staphylococcus aureus* (una causa de intoxicación alimentaria, así como infecciones de la piel, infecciones de la sangre, neumonía, osteomielitis, artritis y abscesos cerebrales) se detecta en cantidades diminutas por la absorción selectiva de la toxina a través de resinas de intercambio iónico o de aglutinación de látex pasiva usando anticuerpos monoclonales. Su pariente, *S. epidermis*, conduce a infecciones de la sangre y contamina los equipos y superficies en hospitales y máquinas y aparatos sanitarios.

Los microorganismos no virales también pueden clasificarse en base a sus propiedades metabólicas tales como la producción de aminoácidos o metabolitos específicos durante las reacciones de fermentación en sustratos tales como glucosa, maltosa o sacarosa. Alternativamente, los microorganismos se pueden tipificar basándose en su sensibilidad a los antibióticos. También se utilizan anticuerpos específicos para antígenos de la superficie celular o proteínas excretadas tales como toxinas para identificar o tipificar microorganismos. Sin embargo, todos los métodos anteriores se basan en el cultivo del microorganismo antes de su posterior ensayo. El cultivo de microorganismos es caro y consume mucho tiempo y también puede verse afectado por contaminación o proliferación de microorganismos menos exigentes. Las técnicas son también relativamente rudimentarias ya que se deben realizar muchas pruebas en la misma muestra con el fin de llegar a un diagnóstico definitivo. La mayoría de los microorganismos no pueden ser cultivados fácilmente en medios conocidos, y por lo tanto, caen por debajo de los niveles de detección cuando una población mixta típica de las diferentes especies de microorganismo está presente en la naturaleza o asociada con organismos superiores.

Otros métodos para la detección e identificación de microorganismos patógenos se basa en el enfoque serológico en el que se producen anticuerpos en respuesta a la infección por el microorganismo. Los meningococos, por ejemplo, pueden clasificarse sobre la base de las diferencias estructurales en sus polisacáridos capsulares. Estos tienen antigenidades diferentes, que permiten la determinación de cinco serogrupos principales, (A, B, C, Y y W-135). Los ensayos inmunoabsorbentes ligados a enzima (ELISA) o el radioinmunoanálisis (RIA) pueden evaluar la producción de tales anticuerpos. Ambos métodos detectan la presencia de anticuerpos específicos producidos por el animal anfitrión durante el curso de la infección. Estos métodos tienen el inconveniente de que se necesita algo de tiempo para que el animal anfitrión produzca un anticuerpo, de este modo a menudo se pasan por alto infecciones muy tempranas. Además, el uso de dichos análisis no es capaz de diferenciar de manera fiable entre una infección pasada y una activa.

Más recientemente, ha habido mucho interés en el uso de métodos moleculares para el diagnóstico de enfermedades infecciosas. Estos métodos ofrecen la detección sensible y específica de microorganismos patógenos. Los ejemplos de tales métodos incluyen el sistema de detección de señales de "ADN ramificado". Este método es un ejemplo que utiliza el principio de hibridación sándwich (Urdea MS et al. Branched DNA amplification

multimers for the sensitive, direct detection of human HIV and hepatitis viruses. Nucleic Acids Symp Ser. 1991; (24):197-200).

Otro método para la detección y clasificación de las bacterias es la amplificación de secuencias de ARN ribosómico 16S. Se ha informado que el ARNr 16S es una diana adecuada para su uso en análisis de amplificación mediante PCR para la detección de especies bacterianas en una variedad de muestras clínicas o ambientales, y con frecuencia se ha utilizado para identificar varios microorganismos específicos debido a que los genes del ARNr 16S muestran polimorfismos específicos de la especie (Cloud, J. L., H. Neal, R. Rosenberry, C. Y. Turenne, M. Jama, D. R. Hillyard, y K. C. Carroll. 2002. J. Clin. Microbiol. 40:400-406). Sin embargo, se requieren cultivos puros de bacterias y después de la amplificación mediante PCR la muestra todavía tiene que ser secuenciada o hibridada a un dispositivo de tipo micro-matriz para determinar la especie (Fukushima M, Kakinuma K, Hayashi H, Nagai H, Ito K; Kawaguchi R. J Clin Microbiol. 2003 Jun; 41(6):2605-15). Tales métodos son costosos, requieren tiempo y trabajo intensivo.

Los autores de la presente invención han desarrollado nuevos métodos para la detección de microorganismos que pueden ser adaptados a la detección general o análisis iniciales de escrutinio para cualquier especie microbiana.

## 15 Descripción de la invención

En un aspecto general, la presente invención se refiere a la reducción de la complejidad de la determinación básica de un genoma microbiano o de un ácido nucleico mediante el tratamiento del ácido nucleico microbiano con un agente que modifica la citosina y la amplificación del ácido nucleico tratado para producir una forma simplificada del genoma o ácido nucleico.

20 En un primer aspecto, la presente invención proporciona un procedimiento para la simplificación de un genoma microbiano o de un ácido nucleico microbiano que comprende:

tratar el genoma o el ácido nucleico microbianos con un agente que modifica la citosina para formar un ácido nucleico microbiano derivado, y

25 amplificar el ácido nucleico microbiano derivado para producir una forma simplificada del genoma o del ácido nucleico microbianos.

En un segundo aspecto, la presente invención proporciona un método para producir una molécula de ácido nucleico específica microbiana que comprende:

tratar una muestra que contiene ADN derivado microbiano con un agente que modifica la citosina para formar ácido nucleico microbiano derivado, y

30 amplificar al menos una parte del ácido nucleico microbiano derivado para formar una molécula de ácido nucleico simplificada que tiene un número total reducido de citosinas en comparación con el correspondiente ácido nucleico microbiano sin tratamiento, en donde la molécula de ácido nucleico simplificada incluye una secuencia de ácido nucleico específica para un microorganismo o un tipo de microorganismo.

35 En un tercer aspecto, la presente descripción proporciona un método para producir una molécula ácido nucleico específica microbiana que comprende:

obtener una secuencia de ADN de un microorganismo;

formar una forma simplificada de la secuencia de ADN microbiano mediante la realización de una conversión de la secuencia de ADN microbiano, cambiando cada citosina a timina de manera que la forma simplificada del ADN microbiano comprenda sustancialmente las bases adenina, guanina y timina; y

40 seleccionar una molécula de ácido nucleico específica microbiana a partir de la forma simplificada del ADN microbiano.

En un cuarto aspecto, la descripción proporciona una molécula de ácido nucleico específica microbiana obtenida mediante el método de acuerdo con el tercer aspecto de la presente invención.

45 En un quinto aspecto, la descripción proporciona el uso del método de acuerdo con el tercer aspecto de la presente invención para obtener sondas o cebadores para unir o amplificar la molécula de ácido nucleico específica microbiana en un ensayo o análisis.

En un sexto aspecto, la descripción proporciona sondas o cebadores obtenidos mediante el quinto aspecto de la presente invención.

La presente invención proporciona un método para detectar la presencia de un microorganismo en una muestra, que comprende:

obtener ADN microbiano a partir de una muestra que se sospecha que contiene el microorganismo;

5 tratar el ácido nucleico microbiano con un agente que modifica la citosina para formar un ácido nucleico microbiano derivado;

proporcionar cebadores capaces de permitir la amplificación de una molécula de ácido nucleico microbiana específica deseada para el ácido nucleico microbiano derivado;

llevar a cabo una reacción de amplificación del ácido nucleico microbiano derivado para formar un ácido nucleico simplificado, y

10 analizar la presencia de un producto de ácido nucleico amplificado que contiene la molécula de ácido nucleico microbiana específica deseada, en donde la detección de la molécula de ácido nucleico microbiana específica deseada es indicativa de la presencia del microorganismo en la muestra.

Si el genoma o el ácido nucleico microbiano es ADN, éste se puede tratar para formar un ADN derivado que se amplifica a continuación para formar una forma simplificada de ADN.

15 Si el genoma o el ácido nucleico microbiano es ARN, éste se puede convertir en ADN antes del tratamiento del genoma o del ácido nucleico microbiano. Alternativamente, el ARN microbiano se puede tratar para producir una molécula de ARN derivada que después se convierte en una molécula de ADN derivada antes de la amplificación. Los métodos de conversión de ARN en ADN son bien conocidos e incluyen el uso de la transcriptasa inversa para formar un ADNc.

20 El genoma o ácido nucleico microbianos se puede obtener a partir del fago, virus, viroides, bacterias, hongos, algas, protozoos, espiroquetas, o organismo unicelular.

El genoma o ácido nucleico microbianos se pueden seleccionar entre ácido nucleico que codifica una proteína, ácido nucleico que no codifica una, las regiones de genes ribosomales de procariotas o de microorganismos eucarióticos unicelulares. Preferiblemente, las regiones de genes ribosomales son 16S o 23S en procariotas y 18S y 28S en el caso de los microorganismos eucarióticos unicelulares. El agente se puede seleccionar entre bisulfito, acetato o citrato. Preferiblemente, el agente es bisulfito de sodio.

25 Preferiblemente, el agente modifica una citosina a un uracilo en cada hebra del ADN genómico microbiano de doble hebra complementario formando dos moléculas de ácido nucleico microbiano derivadas, pero no complementarias. En una forma preferida, la citosina no está metilada como se encuentra típicamente en el ácido nucleico microbiano.

30 Preferiblemente, el ácido nucleico microbiano derivado tiene un número total reducido de citosinas en comparación con el genoma o ácido nucleico microbianos no tratados correspondientes.

Preferiblemente, la forma simplificada del genoma o ácido nucleico microbianos tiene un número total reducido de citosinas en comparación con el genoma o ácido nucleico microbianos no tratados correspondientes.

35 En una forma preferida, el ácido nucleico microbiano derivado contiene sustancialmente bases adenina (A), guanina (G), timina (T) y uracilo (U) y tiene sustancialmente el mismo número total de bases que el genoma o ácido nucleico microbianos no tratados correspondientes.

En otra forma preferida, la forma simplificada del genoma o ácido nucleico microbianos se compone sustancialmente de las bases adenina (A), guanina (G) y timina (T).

40 Preferiblemente, la amplificación se lleva a cabo mediante cualquier medio adecuado tal como reacción en cadena de la polimerasa (PCR), amplificación isotérmica, o amplificación de la señal.

El método de acuerdo con el segundo aspecto de la presente invención puede comprender adicionalmente:

detectar la molécula de ácido nucleico específica microbiana.

En una forma preferida, la molécula de ácido nucleico específica microbiana se detecta:

45 proporcionando un ligando detector capaz de unirse a una región diana de la molécula de ácido nucleico específica microbiana y dejando tiempo suficiente para que el ligando detector se una a la región diana e; y

midiendo la unión del ligando detector a la región diana para detectar la presencia de la molécula de ácido nucleico específica microbiana.

En otra forma preferida, la molécula de ácido nucleico específica microbiana se detecta mediante la separación de un producto de amplificación y la visualización del producto separado. Preferiblemente, el producto de amplificación se separa mediante electroforesis y se detecta mediante la visualización de una o más bandas en un gel.

5 Preferiblemente, la molécula de ácido nucleico específica microbiana no se produce naturalmente en el microorganismo.

En una forma preferida, la molécula de ácido nucleico específica microbiana tiene una secuencia de ácido nucleico indicativa del nivel taxonómico del microorganismo. El nivel taxonómico del microorganismo incluye, pero no está limitando a, familia, género, especie, cepa, tipo o diferentes poblaciones de las mismas o diferentes poblaciones geográficas o bentónicas.

10 En una forma preferida del método de acuerdo con tercer aspecto de la presente descripción, se obtienen las formas simplificadas de dos o más secuencias de ADN microbiano y se comparan las dos o más secuencias para obtener al menos una molécula de ácido nucleico específica microbiana.

En una forma preferida de la presente invención, las moléculas de ácido nucleico se detectan:

15 proporcionando un ligando detector capaz de unirse a una región de la molécula de ácido nucleico y permitiendo un tiempo suficiente para que el ligando detector se una a la región; y

midiendo la unión del ligando detector a la molécula de ácido nucleico para detectar la presencia de la molécula de ácido nucleico.

En otra forma preferida, las moléculas de ácidos nucleicos se detectan mediante la separación de un producto de amplificación y la visualización del producto separado.

20 En situaciones en las que el microorganismo no tiene un genoma de ADN o el genoma o ácido nucleico microbianos son ARN, por ejemplo, un virus ARN, el ARN del genoma viral puede ser convertido primero en ADNc con el fin de tratar el ADN con el agente. El ARN también se puede tratar y el ARN derivado se convierte en ADN antes de la amplificación.

25 Preferiblemente, el ácido nucleico derivado contiene sustancialmente las bases adenina (A), guanina (G), timina (T) y uracilo (U) y tiene sustancialmente el mismo número total de bases que el ácido nucleico microbiano no modificado correspondiente. Es importante destacar que la molécula de ácido nucleico derivado no contiene sustancialmente citosina (C), con la condición de que el ADN microbiano no esté metilado en ninguna de las citosinas.

30 Preferiblemente, el ácido nucleico derivado amplificado contiene sustancialmente las bases A, T y G y tiene sustancialmente el mismo número total de bases que el correspondiente ácido nucleico derivado (y ácido nucleico microbiano no modificado). El ácido nucleico derivado amplificado se denomina ácido nucleico simplificado.

35 En una forma preferida, la molécula de ácido nucleico específica microbiana tiene una secuencia de ácido nucleico indicativa del nivel taxonómico del microorganismo. El nivel taxonómico del microorganismo puede incluir familia, género, especie, cepa, tipo o diferentes poblaciones de las mismas o diferentes poblaciones geográficas o bentónicas. En el caso de las bacterias los autores de la presente invención se pueden adherir al esquema generalmente reconocido, por ejemplo *Bacteria*; *Proteobacteria*; *Betaproteobacteria*; *Neisseriales*; *Neisseriaceae*; *Neisseria*. Las diferentes poblaciones pueden ser polimórficas para cambios de un único nucleótido o una variación que existe en moléculas de ADN que existen en una forma intracelular dentro de un microorganismo (plásmidos o fagémidos), o regiones cromosómicas polimórficas de los genomas de microorganismos tales como islas de patogenicidad.

40 La presente invención también se puede utilizar para reconocer la fluidez de los genomas microbianos y virales, y se puede utilizar para reconocer la naturaleza quimérica de los genomas virales, que pueden estar en piezas independientes, y por lo tanto surgen cepas recién derivadas de la redistribución de regiones genómicas de diferentes animales, por ejemplo, nuevas cepas de la gripe humana como quimeras de segmentos que son recogidas a partir de otros genomas virales de mamífero o de ave.

45 Se apreciará que el método puede llevarse a cabo *in silico* a partir de secuencias conocidas de ácidos nucleicos de los microorganismos en los que una o más citosinas en las secuencias originales se convierte en timina para obtener el ácido nucleico simplificado. La identidad de secuencia se puede determinar a partir de las secuencias convertidas. Tal método *in silico* imita las etapas de tratamiento y amplificación.

50 Cuando una molécula de ácido nucleico específica microbiana ha sido obtenida para cualquier microorganismo dado mediante este método, las sondas o cebadores pueden ser diseñados para asegurar la amplificación de la región de interés en una reacción de amplificación. Por lo tanto, cuando se hayan diseñado las sondas o cebadores, será posible llevar a cabo análisis clínicos o científicos sobre las muestras para detectar un microorganismo dado a un nivel taxonómico dado.

La molécula de ácido nucleico específica microbiana puede ser única o tener un alto grado de similitud dentro de un nivel taxonómico. Una ventaja de la presente invención es la capacidad para simplificar en gran medida las diferencias de bases potenciales entre, o dentro de, los niveles taxonómicos, por ejemplo, de un microorganismo, a una molécula única o a moléculas que tienen una íntima similitud de secuencia. Se pueden utilizar cebadores específicos o un número reducido de cebadores degenerados para amplificar la molécula de ácido nucleico específica microbiana en una muestra dada.

Para el ADN de doble hebra que contiene citosinas, la etapa de tratamiento da como resultado dos ácidos nucleicos derivados (uno para cada hebra complementaria), que contienen cada uno las bases adenina, timina, guanina y uracilo. Los dos ácidos nucleicos derivados se producen a partir de las dos hebras sencillas del ADN de doble hebra. Los dos ácidos nucleicos derivados preferiblemente no tienen citosinas pero todavía tienen el mismo número total de bases y la longitud de secuencia que la molécula de ADN original no tratada. Es importante destacar que los dos ácidos nucleicos derivados no son complementarios entre sí y forman un molde de la hebra superior en inferior para la amplificación. Se pueden utilizar una o más de las hebras como diana para la amplificación para producir la molécula de ácido nucleico simplificada. Durante la amplificación de los ácidos nucleicos derivados, los uracilos en la parte superior (o hebra inferior) se sustituyen por timinas en la correspondiente forma simplificada amplificada de los ácidos nucleicos. A medida que continúa la amplificación, la parte superior (y/o hebra inferior si se amplifica) se diluirá a medida que cada nueva hebra complementaria tenga solamente las bases adenina, guanina, timina.

Se apreciará que este aspecto de la invención también incluye moléculas de ácido nucleico que tienen secuencias complementarias a la molécula de ácido nucleico específica microbiana, y moléculas de ácidos nucleicos que pueden hibridarse, preferiblemente bajo condiciones restrictivas, con la molécula de ácido nucleico específica microbiana.

En la presente invención se pueden usar sondas o cebadores que son indicativos de tipos representativos de microorganismo que se pueden utilizar para determinar si cualquier microorganismo está presente en una muestra dada. Se pueden usar otras sondas específicas del tipo microbiano para detectar o identificar realmente un tipo, subtipo, variante y genotipo dados de ejemplos de microorganismo.

Cuando una molécula de ácido nucleico específica microbiana se ha obtenido o identificado para cualquier microorganismo dado, se pueden diseñar sondas o cebadores para asegurar la amplificación de la región de interés en una reacción de amplificación. Es importante señalar que ambas hebras de un genoma tratado y convertido de este modo, (en lo sucesivo denominado "ácido nucleico derivado") pueden ser analizadas para el diseño de cebadores, ya que el tratamiento o la conversión conducen a asimetrías de la secuencia, y por lo tanto se requiere diferentes secuencias de cebadores para la detección de las hebras "superior" e "inferior" del mismo locus, (también conocidas como hebras de "Watson" y "Crick"). Por lo tanto, hay dos poblaciones de moléculas, el genoma convertido, tal como existe inmediatamente después de la conversión, y la población de moléculas que resulta después de que el ácido nucleico derivado sea replicado por medios enzimológicos convencionales (PCR) o por métodos tales como amplificación isotérmica. Los cebadores se diseñan típicamente para la hebra superior convertida por conveniencia pero también se pueden generar los cebadores para la hebra inferior. De este modo, será posible llevar a cabo análisis clínicos o científicos sobre las muestras para detectar un microorganismo dado.

Los cebadores o sondas se pueden diseñar para permitir la amplificación de regiones específicas del ácido nucleico derivado. En una forma preferida, los cebadores producen la amplificación de la molécula de ácido nucleico específica microbiana.

Se describen un kit para detectar una molécula de ácido nucleico específica microbiana que comprende los cebadores o sondas de acuerdo con quinto aspecto de la presente descripción junto con uno o más reactivos o componentes para una reacción de amplificación.

Preferiblemente, el microorganismo se selecciona del fago, virus, viroides, bacterias, hongos, algas, protozoos, espiroquetas, organismo unicelular, o cualquier otro microorganismo, por variada que sea su clasificación, tal como el Reino Protista de Margulis, L., et al 1990, Handbook of Protocista, Jones and Bartlett, Publishers, Boston USA, o microorganismos que están asociados con los seres humanos, tal como se define en Harrison's Principles of Internal Medicine, 12<sup>a</sup> Edición, editado por J D Wilson et al., McGraw Hill Inc, así como ediciones posteriores. También incluye todos los microorganismos descritos en relación con las condiciones humanas definidas en OMIM, Online Mendelian Inheritance in Man, [www.ncbi.gov](http://www.ncbi.gov).

El microorganismo puede ser un patógeno, una muestra ambiental de origen natural, microorganismo acuático o aéreo, (o un organismo existente o que sea transportado en un medio líquido o gaseoso), ya sea maduro o en forma de esporas, o bien extracelularmente o intracelularmente, o asociado con un forma de vida quimérica, o que existe como ectocomensal entre dos o más formas de vida, tales como un microbio asociado con un líquen, o un microbio asociado con una célula bacteriana.

Es posible analizar la presencia de virus o viroides ARN convirtiendo en primer lugar su genoma de ARN en una forma de ADNc a través de la transcripción inversa y modificando después el ADNc por medio del reactivo. Esto

supera el problema de cualquier metilación existente en las citosinas en los virus ARN, ya que la transcriptasa inversa copiará éstas como si fueran citosinas regulares.

Preferiblemente, el agente modifica la citosina no metilada a uracilo que a continuación se reemplaza como una timina durante la amplificación del ácido nucleico derivado. Preferiblemente, el agente utilizado para modificar la citosina es bisulfito de sodio. También se pueden utilizar otros agentes que modifican de manera similar la citosina no metilada, pero no la citosina metilada en el método de la invención. Los ejemplos incluyen, pero no están limitados a bisulfito, acetato o citrato. Preferiblemente, el agente es bisulfito de sodio, un reactivo, que en presencia de agua, modifica la citosina a uracilo.

El bisulfito de sodio ( $\text{NaHSO}_3$ ) reacciona fácilmente con el enlace doble 5,6 de la citosina para formar un intermedio de reacción de citosina sulfonado que es susceptible de desaminación, y en presencia de agua da lugar a un sulfito de uracilo. Si fuera necesario, el grupo sulfito se puede eliminar en condiciones alcalinas suaves, dando como resultado la formación de uracilo. Por lo tanto, potencialmente todas las citosinas se convertirán en uracilos. Cualquier citosina metilada, sin embargo, no puede ser convertida por el reactivo modificador debido a la protección por metilación.

La presente invención se puede adaptar para ayudar a eludir algunos de los problemas emergentes que se ponen de manifiesto por la enorme variación genómica inesperada entre los productos aislados de la misma especie bacteriana, (2005, Tettelin, H., et al., Proc. Natl. Acad. Sci. USA. 102, 13950-13955; Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial "pan-genome"). Todos los productos aislados de esta especie bacteriana tienen un genoma "esencial" de genes que codifica proteínas que representa aproximadamente 80% de la reserva de genes, además de un genoma prescindible que consiste en genes que codifican proteínas específicos de la cepa y parcialmente compartidos. Al tratar el gen o los genes 23S presentes dentro de una población bacteriana mediante los métodos de acuerdo con la presente invención, los autores de la invención pueden abordar un componente codificante no proteico esencial que está presente en todos los productos aislados bacterianos.

La presente invención es adecuada para análisis clínicos, medioambientales, forenses, de guerra bacteriológica o científicos para los microorganismos, en los que es útil la identidad inicial anterior o a nivel de especie, con el fin de determinar primero el grupo general al cual pertenece el organismo. Los ejemplos incluyen, pero no están limitados a, el diagnóstico de la enfermedad en cualquier organismo, (ya sea vertebrados, invertebrados, procariotas o eucariotas, por ejemplo, enfermedades de plantas y animales, enfermedades de las fuentes de alimento humano, tales como las piscifactorías y los criaderos de ostras), el escrutinio o el muestreo de fuentes ambientales ya sean naturales o contaminadas, la determinación de la contaminación de cultivos celulares o de huevos fertilizados *in vitro* para la producción de blastocistos humanos en clínicas de fertilización *in vitro* o para la cría de animales. Tiene especial trascendencia la detección de microorganismos en entornos forenses o en contextos de guerra biológica.

A lo largo de esta memoria descriptiva, a menos que el contexto requiera otra cosa, se entenderá que la palabra "comprender", o variaciones tales como "comprende" o "que comprende", implica la inclusión de un elemento declarado, número entero o etapa, o grupo de elementos, números enteros o etapas, pero no la exclusión de cualquier otro elemento, número entero o etapa, o grupo de elementos, números enteros o etapas.

Cualquier discusión de documentos, actos, materiales, dispositivos, artículos o similares que ha sido incluida en la presente memoria tiene únicamente el propósito de proporcionar un contexto para la presente invención. No se debe tomar como una admisión de que cualquiera o todos estos asuntos forman parte de la base de la técnica anterior o eran el conocimiento general común en el campo relevante para la presente invención que existía en Australia antes del desarrollo de la presente invención.

A fin de que la presente invención pueda entenderse más claramente, se describirán las formas de realización preferidas con referencia a los dibujos y ejemplos siguientes.

#### 45 Breve descripción de los dibujos

La Figura 1 muestra el alineamiento de parte del gen *iga* de *Neisseria meningitidis* y *Neisseria gonorrhoeae* antes y después de la simplificación genómica. Como puede verse, antes de la simplificación genómica, se requerirían un total de 512 combinaciones de sondas para la detección universal de especies de *Neisseria* (74% de similitud de secuencia), en comparación con sólo 2 combinaciones después de la simplificación para formar ácido nucleico derivado (97% de similitud de secuencia). (El SEQ ID NO aparece en la lista después de cada secuencia).

La Figura 2 muestra el uso de sondas de INA para aumentar aún más la similitud de secuencia de las secuencias simplificadas, ya que las sondas de INA pueden tener de longitud más corta que las sondas de oligonucleótidos convencionales. La combinación del procedimiento de simplificación genómico con sondas de INA permite la selección y el uso de sondas con 100% de similitud de secuencia con la secuencia diana. (El SEQ ID NO aparece en la lista después de cada secuencia).

- 5 La Figura 3 muestra la simplificación genómica para diferenciar entre especies estrechamente relacionadas usando alineamientos del gen iga de *Neisseria* y *Haemophilus*. Como puede verse, el método de la presente invención permite la simplificación del material genómico con el fin de producir sondas específicas de especie. Además, a pesar de la simplificación del ADN genómico, todavía se permite la diferenciación entre *Neisseria* y las especies de *Haemophilus* estrechamente relacionadas. (El SEQ ID NO aparece en la lista después de cada secuencia).
- 10 La Figura 4 muestra el alineamiento del gen tuf estreptocócico antes y después de la simplificación genómica en 10 especies diferentes de estreptococos. Antes del tratamiento, se requerirían un total de 12.288 combinaciones de sondas que para el cebador universal del gen tuf. Después de simplificación genómica, solamente se requerirían 64 combinaciones de sondas para la detección universal. Además, la similitud de secuencia antes de la simplificación es sólo de 67,5%, que aumenta a 85% después de la simplificación. (El SEQ ID NO aparece en la lista después de cada secuencia).
- 15 La Figura 5 muestra el alineamiento de los genes de la enterotoxina estafilocócica antes y después de la simplificación genómica. Antes de tratamiento con bisulfito, se requerirían un total de 1.536 combinaciones de sondas para el cebador universal del gen de la enterotoxina estafilocócica. Después de la simplificación genómica sólo se requerirían 64 combinaciones de sondas para la detección universal. (El SEQ ID NO aparece en la lista después de cada secuencia).
- 20 La Figura 6 muestra el alineamiento del gen de neuraminidasa de Influenza del grupo A y B de diferentes cepas de influenza antes y después de la simplificación genómica. Antes del tratamiento, se requerirían un total de 2.048 combinaciones de sondas para el cebador universal del gen de la neuraminidasa del grupo A y B. Después de simplificación genómica sólo se requerirían 48 combinaciones de sonda para la detección universal. Además, la similitud de secuencia antes de la simplificación es sólo de 50%, que aumenta a 75% después de la simplificación. (El SEQ ID NO aparece en la lista después de cada secuencia).
- 25 La Figura 7 muestra el alineamiento del gen VP4 de rotavirus antes y después de la simplificación genómica. Antes del tratamiento, se requerirían un total de 512 combinaciones de sondas para el cebador universal del gen VP4 de rotavirus. Después de simplificación genómica sólo se requerirían 32 combinaciones de sonda para la detección universal. (El SEQ ID NO aparece en la lista después de cada secuencia).
- 30 La Figura 8 muestra los productos de amplificación obtenidos mediante PCR a partir de las regiones del gen ribosomal 23S genómicamente simplificadas de bacterias Gram positivas y Gram negativas, detectándose amplicones adecuados como bandas de longitud específica mediante electroforesis en gel de agarosa. La flecha indica el tamaño esperado de los amplicones con relación a los marcadores de tamaño convencionales ejecutados en la calle del marcador, (M). El uso de cebadores específicos para bacterias Gram negativas revela bandas sólo en las seis calles Gram negativas (panel superior). El uso de cebadores específicos para las bacterias Gram positivas revela sólo bandas en las seis calles Gram positivas (panel inferior).
- 35 La Figura 9 muestra los productos de amplificación obtenidos mediante PCR a partir de las regiones del gen ribosomal 23S genómicamente simplificadas de *E. coli* (Calle 1) y *K. pneumoniae*, (Calle 3). La especificidad de la amplificación se ilustra por la ausencia de productos de amplificación a partir de las 10 especies de bacterias restantes.
- 40 La Figura 10 muestra el producto de amplificación obtenido mediante PCR a partir de las regiones del gen ribosomal 23S genómicamente simplificadas usando cebadores específicos para *Neisseria*.
- La Figura 11 muestra el producto de amplificación obtenido mediante PCR a partir de un gen codificante de la proteína de la región genómicamente simplificada del gen recA de *E. coli*. La especificidad del amplicón se ilustra por la presencia del amplicón recA de *E. coli* y su ausencia de las otras 11 especies de bacterias.
- 45 La Figura 12 muestra los productos de amplificación obtenidos mediante PCR a partir de las regiones del gen ribosomal 23S genómicamente simplificadas usando cebadores específicos para estafilococos.
- La Figura 13 muestra los productos de amplificación obtenidos mediante PCR a partir de las regiones del gen ribosomal 23S genómicamente simplificadas usando cebadores específicos para estreptococos.
- 50 La Figura 14 muestra los productos de amplificación obtenidos mediante PCR de un gen codificante de la proteína de la región genómicamente simplificada del gen recA de *Staphylococcus epidermidis*. Las dos bandas (flechas) representan el traslado de amplicones de la primera ronda, (banda superior) y la segunda ronda (banda inferior), amplificaciones de PCR.
- La Figura 15 muestra la detección de amplicones utilizando cebadores específicos dirigidos a los genes ribosomales 23S genómicamente simplificados de *Chlamydia trachomatis*.

La Figura 16 muestra las secuencias de secuencias de ADNr 23S genómicas normales y genómicamente simplificadas de *Staphylococcus epidermidis*. (El SEQ ID NO aparece en la lista después de cada secuencia).

La Figura 17 muestra las secuencias de secuencias genómicas y genómicamente simplificadas del gen recA de *E. coli*. (El SEQ ID NO aparece en la lista después de cada secuencia).

## 5 Modo de llevar a cabo la invención

### Definiciones

El término "simplificación genómica" según se utiliza en la presente memoria significa que el ácido nucleico genómico (o de otro tipo) se modifica pasando de comprender cuatro bases adenina (A), guanina (G), timina (T) y citosina (C) a contener sustancialmente las bases adenina (A), guanina (G), timina (T), pero teniendo todavía sustancialmente el mismo número total de bases.

El término "derivado de ácido nucleico" según se utiliza en la presente memoria significa un ácido nucleico que contiene sustancialmente las bases A, G, T y U (o alguna otra base o entidad de tipo base distinta de A, G o T) y tiene sustancialmente el mismo total número de bases que el correspondiente ácido nucleico microbiano no modificado. Sustancialmente todas las citosinas en el ADN microbiano se habrán convertido en uracilo durante el tratamiento con el agente. Se apreciará que las citosinas alteradas, por ejemplo mediante metilación, no necesariamente pueden ser convertidas en uracilo (o alguna otra base o entidad de tipo base distinta de A, G o T). Como el ácido nucleico microbiano no contiene típicamente citosina metilada (u otras alteraciones de citosina) la etapa tratada convierte preferiblemente todas las citosinas. Preferiblemente, la citosina es modificada a uracilo.

El término "ácido nucleico simplificado" según se utiliza en la presente memoria, significa el producto resultante de ácido nucleico obtenido después de la amplificación del ácido nucleico derivado. El uracilo en el ácido nucleico derivado se reemplaza a continuación como una timina (T) durante la amplificación del ácido nucleico derivado para formar la molécula de ácido nucleico simplificada. El producto resultante tiene sustancialmente el mismo número de bases totales que el correspondiente ácido nucleico microbiano no modificado pero está sustancialmente formado por una combinación de tres bases (A, G y T).

El término "secuencia simplificada" según se utiliza en la presente memoria significa la secuencia de ácido nucleico resultante obtenida después de la amplificación del ácido nucleico derivado para formar un ácido nucleico simplificado. La secuencia simplificada resultante tiene sustancialmente el mismo número de bases totales que la correspondiente secuencia de ácido nucleico microbiano no modificada pero está formada sustancialmente por una combinación de tres bases (A, G y T).

El término "secuencia no convertida" según se utiliza en la presente memoria significa la secuencia de ácido nucleico del ácido nucleico microbiano antes del tratamiento y la amplificación. Una secuencia no convertida es típicamente la secuencia del ácido nucleico microbiano de origen natural.

El término "modifica" según se utiliza en la presente memoria significa la conversión de una citosina en otro nucleótido. Preferiblemente, el agente modifica la citosina no metilada a uracilo para formar un ácido nucleico derivado.

El término "agente que modifica la citosina" según se utiliza en la presente memoria significa un agente que es capaz de convertir la citosina en otra entidad química. Preferiblemente, el agente modifica la citosina a uracilo que luego se reemplaza como una timina durante la amplificación del ácido nucleico derivado. Preferiblemente, el agente utilizado para modificar la citosina es bisulfito de sodio. También se pueden utilizar otros agentes que modifican de manera similar la citosina, pero no la citosina metilada en el método de la invención. Los ejemplos incluyen, pero no están limitados a bisulfito, acetato o citrato. Preferiblemente, el agente es bisulfito de sodio, un reactivo, que en presencia de condiciones ácidas acuosas, modifica la citosina a uracilo. El bisulfito de sodio ( $\text{NaHSO}_3$ ) reacciona fácilmente con el enlace doble 5,6 de la citosina para formar un intermedio de reacción de citosina sulfonada que es susceptible de desaminación, y en presencia de agua da lugar a un sulfito de uracilo. Si fuera necesario, el grupo sulfito puede ser eliminado en condiciones alcalinas suaves, dando como resultado la formación de uracilo. Por lo tanto, potencialmente todas las citosinas se convertirán en uracilos. Cualquier citosina metilada, sin embargo, puede no ser convertida por el reactivo modificador debido a la protección mediante metilación. Se apreciará que la citosina (o cualquier otra base) podría ser modificada mediante medios enzimáticos para lograr un ácido nucleico derivado tal como se ilustra en la presente invención.

Existen dos métodos genéricos amplios por medio de los cuales se pueden modificar las bases de los ácidos nucleicos: químicos y enzimáticos. Por lo tanto, la modificación de la presente invención también puede llevarse a cabo por enzimas de origen natural, o por enzimas construidas o seleccionadas artificialmente aún por referir. El tratamiento químico, tal como las metodologías con bisulfito, pueden convertir citosina en uracilo mediante etapas químicas adecuadas. De manera similar, las citosina desaminasas, por ejemplo, pueden llevar a cabo una conversión para formar un ácido nucleico derivado. El primer informe sobre citosina desaminasas en el conocimiento

de los autores de la presente invención es 1932, Schmidt, G., Z. Physiol. Chem., 208, 185; (véanse también 1950, Wang, T.P., Sable, H.Z., Lampen, J.O., J. Biol. Chem, 184, 17-28, Enzymatic deamination of cytosines nucleosides). En este primer trabajo, no se obtuvo citosina desaminasa libre de otras nucleodesaminasas, sin embargo, Wang et al. fueron capaces de purificar dicha actividad de la levadura y *E. coli*. Así, cualquier conversión enzimática de la citosina para formar un ácido nucleico derivado que finalmente da como resultado la inserción de una base durante la siguiente replicación en esa posición, que es diferente de una citosina, producirá un genoma simplificado. La conversión química y enzimática para proporcionar un derivado seguido por un genoma simplificado es aplicable a cualquier nucleobase, ya se trate de purinas o pirimidinas en ácidos nucleicos origen natural de microorganismos.

El término "forma simplificada del genoma o ácido nucleico", según se utiliza en la presente memoria significa que un genoma o ácido nucleico, ya sea de origen natural o sintético, que por lo general contiene las cuatro bases comunes G, A, T y C, ahora se compone principalmente de sólo tres bases, G, A y T ya que la mayoría o la totalidad de las C en el genoma se han convertido en T mediante la modificación química adecuada y procedimientos posteriores de amplificación. La forma simplificada del genoma significa que la complejidad genómica relativa se reduce de un fundamento de cuatro bases a una composición de tres bases.

El término «entidad de tipo base» según se utiliza en la presente memoria, significa una entidad que se forma por la modificación de la citosina. Una entidad de tipo base puede ser reconocida por una ADN polimerasa durante la amplificación de un ácido nucleico derivado y la polimerasa hace que A, G o T se coloquen sobre la hebra ADN complementaria recién formada en la posición opuesta a la entidad de tipo base en el ácido nucleico derivado. Típicamente, la entidad de tipo base es el uracilo que ha sido modificado a partir de la citosina al correspondiente ácido nucleico microbiano no tratado. Los ejemplos de una entidad de tipo base incluye cualquier nucleobase, ya sea purina o pirimidina.

El término "reducción de la complejidad relativa", según se utiliza en la presente memoria, hace referencia a la longitud de la sonda, a saber, el aumento de la longitud media de la sonda que se requiere para lograr la misma especificidad y nivel de hibridación de una sonda a un locus específico, en un conjunto dado de condiciones moleculares en dos genomas del mismo tamaño, donde el primer genoma es "tal cual" y consiste en las cuatro bases, G, A, T y C, mientras que el segundo genoma es exactamente de la misma longitud, pero algunas citosinas, (idealmente todas las citosinas), se han convertido en timinas. El locus sometido a ensayo está en la misma ubicación en el genoma original no convertido, así como el genoma convertido. Por término medio, una sonda de 11 unidades tendrá una ubicación única a la que se hibridará perfectamente en un genoma regular de 4.194.304 bases consistentes en las cuatro bases G, A, T y C, ( $4^{11}$  es igual a 4.194.304). Sin embargo, una vez que dicho genoma regular de 4.194.304 bases ha sido convertido por el bisulfito u otros medios adecuados, este genoma convertido se compone ahora solamente de tres bases y es claramente menos complejo. Sin embargo, la consecuencia de esta *disminución* de la complejidad genómica es que la sonda de los autores de la presente invención previamente de 11 unidades única ya no tiene un sitio único al que se puede hibridar dentro del genoma simplificado. Ahora hay muchas otras posibles localizaciones equivalentes de secuencias de 11 bases que se han presentado *de novo* como consecuencia de la conversión con bisulfito. Ahora se requerirá una sonda de 14 unidades para encontrar e hibridarse con el locus original. Aunque en un principio puede parecer contrario al sentido común, se requiere de este modo un *aumento* de la longitud de la sonda para detectar la ubicación original en lo que es ahora un genoma de tres bases simplificado, puesto que más genoma parece el mismo, (tiene secuencias más similares). Así, la reducida *complejidad genómica relativa*, (o simplicidad del genoma de tres bases), significa que se tienen que diseñar sondas *más* largas para encontrar el único sitio original.

El término "reducción de la complejidad genómica relativa" según se utiliza en la presente memoria se puede medir por medio del aumento de las longitudes de las sonda susceptibles de ser específicas de microbio en comparación con el ADN no modificado. Este término también incorpora el tipo de secuencias de sondas que se utilizan en la determinación de la presencia de un microorganismo. Estas sondas pueden tener cadenas principales no convencionales, tales como las de PNA o LNA o adiciones modificadas a una cadena principal tales como las descritas en el INA. Por lo tanto, se considera que un genoma tiene una *complejidad relativa* reducida, con independencia de si la sonda tiene componentes adicionales tales como pseudonucleótidos intercalantes, tal como en el INA. Los ejemplos incluyen, pero no están limitados a, ADN, ARN, ácido nucleico cerrado (LNA), ácido peptidonucleico (PNA), MNA, ácido nucleico altritol (ANA), ácido nucleico hexitol (HNA), ácido nucleico intercalante (INA), ácido nucleico ciclohexanilo (CNA) y mezclas de los mismos y sus híbridos, así como las modificaciones de átomo de fósforo de los mismos, tales como, pero no limitadas a fosforotioatos, fosfolatos de metilo, fosforamiditas, fosforditatos, fosforoselenoatos, fosfotriésteres y fosforanoatos. Los nucleótidos de origen no natural incluyen, pero no están limitados a los nucleótidos comprendidos dentro de ADN, ARN, PNA, INA, HNA, MNA, ANA, LNA, CNA, CeNA, TNA, (2'-NH)-TNA, (3'-NH)-TNA,  $\alpha$ -L-Ribo-LNA,  $\alpha$ -L-Xilo-LNA,  $\beta$ -D-Xilo-LNA,  $\alpha$ -D-Ribo-LNA, [3,2,1]-LNA, Biciclo-ADN, 6-Amino-Biciclo-ADN, 5-epi-Biciclo-ADN,  $\alpha$ -Biciclo-ADN, ADN-Triciclo, Biciclo[4,3,0]-ADN, Biciclo[3,2,1]-ADN, Biciclo[4,3,0]amida-ADN,  $\beta$ -D-Ribopiranosil-NA,  $\alpha$ -L-Lixopiranosil-NA, 2'-R-ARN,  $\alpha$ -L-ARN o  $\alpha$ -D-ARN,  $\beta$ -D-ARN. Además se pueden usar compuestos que no contienen fósforo para la unión a los nucleótidos tales como, pero no limitados a metiliminometilo, formacetato, tioformacetato y grupos de unión que comprenden amidas. En particular, los ácidos nucleicos y análogos de ácido nucleico pueden comprender uno o más pseudonucleótidos

intercalantes (IPN). La presencia de IPN no es parte de la descripción de la complejidad de las moléculas de ácido nucleico, ni es la parte central de esa complejidad, tal como en PNA.

5 Por "INA" se quiere significar un ácido nucleico intercalante de acuerdo con la enseñanza de los documentos WO 03/051901, WO 03/052132, WO 03/052133 y WO 03/052134 (Unest A/S). Un INA es un oligonucleótido o un análogo de oligonucleótido que comprende una o más moléculas de pseudonucleótidos intercalantes (IPN).

Por "HNA" se quieren significar ácidos nucleicos como describen por ejemplo Van Aetschot et al., 1995.

Por "MNA" se quieren significar ácidos nucleicos como describen Hossain et al, 1998.

"ANA" hace referencia a los ácidos nucleicos descritos por Allert et al, 1999.

10 "LNA" puede ser cualquier molécula de LNA como se describe en el documento WO 99/14226 (Exiqon), Preferiblemente, el LNA se selecciona entre las moléculas descritas en el resumen del documento WO 99/14226. Más preferiblemente, el LNA es un ácido nucleico como describen Singh et al, 1998, Koshkin et al, 1998 u Obika et al., 1997.

"PNA" hace referencia a ácidos peptidonucleicos, como describen Nielsen et al, 1991.

15 "Reducción de la complejidad relativa" según se utiliza en la presente memoria, no hace referencia al orden en el que aparecen las bases, tal como cualquier diferencia de complejidad matemática entre una secuencia que es ATATATATATATAT (SEQ ID NO: 1) frente a una de la misma longitud que es AAAAAAATTTTTTTT (SEQ ID NO: 2), ni tampoco hace referencia a los datos de re-asociación original de los tamaños relativos del genoma, (y por inferencia, las complejidades genómicas), introducidos en las publicaciones científicas por Waring, M. & Britten RJ1966, Science, 154, 791-794; y Britten, RJ y Kohne D E., 1968, Science, 161, 529-540 y referencias anteriores de  
20 las mismas que se derivan de los informes de la Carnegie Institution of Washington Yearbook.

"Complejidad genómica relativa" según se utiliza en la presente memoria hace referencia a una posición inalterada de bases en dos genomas a la que se accede por medio de sondas moleculares (tanto los genomas originales como los no convertidos tienen bases en las posiciones invariantes 1 a n. En el caso del genoma humano haploide de 3 mil millones de pares de bases de una hembra humana concreta, se define que las posiciones invariantes están  
25 entre 1 y n, donde n es 3000000000. Si en la secuencia de 1 a n, la base  $i$  es una C en el genoma original, la base  $i$  es una T en el genoma convertido.

El término "ácido nucleico genómico" según se utiliza en la presente memoria incluye ARN (de procarionota y eucariota unicelular), ADN, ácido nucleico que codifica proteínas, ácido nucleico que no codifica proteínas, y regiones de genes ribosomales de microorganismos procarionóticos y eucarióticos unicelulares microbianos.

30 El término "genoma microbiano" según se utiliza en la presente memoria abarca ácidos nucleicos cromosómicos así como extracromosómicos, así como residentes temporales de ese genoma, tales como plásmidos, bacteriófagos y elementos móviles en el sentido más amplio. El "genoma" tiene un componente esencial como se ilustra mediante *S. galactiae*, así como también tiene posiblemente elementos codificantes y no codificantes que varían entre los distintos productos aislados.

35 El término "ADN microbiano derivado", según se utiliza en la presente memoria incluye ADN obtenido directamente a partir de un microorganismo u obtenido indirectamente mediante la conversión de ARN microbiano en ADN mediante cualquier método conocido o adecuado, tal como la transcriptasa inversa.

40 El término "microorganismo", según se utiliza en la presente memoria incluye fagos, virus, viroides, bacterias, hongos, algas, protozoos, espiroquetas, organismos unicelulares, o cualquier otro microorganismo, por variada que sea su clasificación, tal como el Reino Protista de Margulis, L., et al 1990, Handbook of Protoctista, Jones and Bartlett, Publishers, Boston USA, o microorganismos que están asociados con los seres humanos, tal como se define en Harrison's Principles of Internal Medicine, 12<sup>a</sup> Edición, editado por J D Wilson et al., McGraw Hill Inc, así como ediciones posteriores. También incluye todos los microorganismos descritos en relación con las condiciones humanas definidas en OMIM, Online Mendelian Inheritance in Man, [www.ncbi.gov](http://www.ncbi.gov).

45 El término "molécula de ácido nucleico específica microbiana" según se utiliza en la presente memoria significa una molécula que ha sido determinada u obtenida utilizando el método de acuerdo con la presente invención que tiene una o más secuencias específicas para un microorganismo.

50 El término "nivel taxonómico del microorganismo", según se utiliza en la presente memoria incluye familia, género, especie, cepa, tipo o diferentes poblaciones de las mismas o diferentes poblaciones geográficas o bentónicas. Si bien en el caso de las bacterias se utiliza el esquema generalmente reconocido, por ejemplo *Bacteria*; *Proteobacteria*; *Betaproteobacteria*; *Neisseriales*; *Neisseriaceae*; *Neisseria*. Las diferentes poblaciones pueden ser polimórficas para cambios de un único nucleótido o una variación que existe en moléculas de ADN que existen en una forma intracelular dentro de un microorganismo (plásmidos o fagémidos), o regiones cromosómicas polimórficas

de los genomas de microorganismos tales como islas de patogenicidad. La fluidez de los genomas microbianos y virales, es reconocida, e incluye la naturaleza quimérica de los genomas virales, que pueden estar en piezas independientes de ácido nucleico. Por lo tanto surgen cepas recién derivadas de la redistribución de regiones genómicas de diferentes animales, por ejemplo, nuevas cepas de la gripe humana como quimeras de segmentos que son recogidas a partir de otros genomas virales de mamífero o de ave.

El término "íntima similitud de secuencia " según se utiliza en la presente memoria incluye la definición anterior de complejidad relativa de secuencia y longitudes de sonda como una medida.

### **Materiales y métodos**

#### **Extracción de ADN**

En general, el ADN microbiano (o ARN viral) se puede obtener de cualquier fuente adecuada. Los ejemplos incluyen, pero no están limitados a, cultivos de células, caldos de cultivo, muestras ambientales, muestras clínicas, fluidos corporales, muestras líquidas, muestras sólidas tales como tejido. El ADN microbiano de las muestras se puede obtener mediante procedimientos convencionales. Un ejemplo de una extracción adecuada es el siguiente. La muestra de interés se coloca en 400 µl de hidrocloreuro de guanidinio 7 M, EDTA 5 mM, Tris/HCl 100 mM pH 6,4, Triton-X-100 al 1%, proteinasa K (Sigma) 50 mM, ARNt de levadura de 100 µg/ml. La muestra se homogeneiza a fondo con una mano de mortero desechable de 1,5 ml y se deja durante 48 horas a 60°C. Después de la incubación la muestra se somete a cinco ciclos de congelación/descongelación en hielo seco durante 5 minutos/95°C durante 5 minutos. La muestra se somete después a vórtice y se centrifuga en una microcentrífuga durante 2 minutos para sedimentar los residuos celulares. Se elimina el sobrenadante en un tubo limpio, se diluye para reducir la concentración de sal, después se extrae con fenol:cloroformo, se precipita con etanol y se resuspende en 50 µl de Tris 10 mM/EDTA 0,1 mM.

Específicamente, las extracciones de ADN a partir de bacterias Gram positivas y Gram negativas cultivadas en placas de agar convencionales (con los requerimientos nutricionales específicos para cada especie) se realizaron como sigue.

Para la extracción de ADN a partir de bacterias gram negativas, el protocolo fue el siguiente:

a) Utilizando un palillo de dientes estéril se desprendieron las colonias bacterianas de la placa de cultivo en un tubo de centrifuga estéril de 1,5 ml.

b) Se añadieron 180 µl de tampón de extracción de tiocianato de guanidinio (tiocianato de guanidinio 7 M, EDTA 5 mM (pH 8,0), Tris/HCl 40 mM pH 7,6, Triton-X-100 al 1%) y la muestra se mezcló para resuspender las colonias bacterianas.

c) Se añadieron 20 µl (20 mg/ml) de proteinasa K y las muestras se mezclaron bien.

d) Las muestras se incubaron a 55°C durante 3 horas para lisar las células.

e) Se añadieron 200 µl de agua a cada muestra y las muestras se mezclaron pipeteando suavemente.

f) Se añadieron 400 µl de fenol/cloroformo/alcohol isoamílico (25:24:1) y las muestras se sometieron a vórtice durante 2 x 15 segundos.

g) Las muestras se centrifugaron a continuación en una microcentrífuga a 14.000 rpm durante 4 minutos.

h) La fase acuosa se separó en un tubo de centrifuga limpio de 1,5 ml.

i) Se añadieron 400 µl de fenol/cloroformo/alcohol isoamílico (25:24:1) y las muestras se sometieron a vórtice durante 2 x 15 segundos.

j) Las muestras se centrifugaron a continuación en una microcentrífuga a 14.000 rpm durante 4 minutos.

k) La fase acuosa se separó en un tubo de centrifuga limpio de 1,5 ml.

l) Se añadieron 800 µl de etanol del 100% a cada muestra, la muestra se sometió a vórtice brevemente, después se dejó a -20°C durante 1 hora.

m) Las muestras se centrifugaron en una microcentrífuga a 14.000 rpm durante 4 minutos a 4°C.

n) Los sedimentos de ADN se lavaron con 500 µl de etanol del 70%.

o) Las muestras se centrifugaron en una microcentrífuga a 14.000 rpm durante 5 minutos a 4°C, el etanol se descartó y los sedimentos se secaron al aire durante 5 minutos.

p) Por último, el ADN se resuspendió en 100 µl de Tris/HCl 10 mM pH 8,0, EDTA 1 mM pH 8,0.

q) La concentración y la pureza del ADN se calcularon midiendo la absorbancia de la disolución a 230, 260, 280 nm.

Para la extracción de ADN de las bacterias Gram positivas el protocolo fue el siguiente:

- 5 a) Utilizando un palillo de dientes estéril se desprendieron colonias bacterianas de la placa de cultivo en un tubo de centrifuga estéril de 1,5 ml.
- b) Se añadieron 180 µl de lisozima de 20 mg/ml (Sigma) y 200 µg de lisostafina (Sigma) a cada muestra y las muestras se mezclaron suavemente para resuspender las colonias bacterianas.
- c) Las muestras se incubaron a 37°C durante 30 minutos para degradar la pared celular.
- 10 d) Las muestras se procesaron a continuación y se extrajo el ADN de acuerdo con el protocolo QIAamp DNA Mini Kit para bacterias Gram positivas.

Extracción de ADN de muestras de citología de pacientes.

- a) La muestra se agitó enérgicamente a mano para volver a suspender las células sedimentadas y para garantizar la homogeneidad de la disolución.
- 15 b) Se transfirieron 4 ml de las células resuspendidas a un tubo de centrifuga Costar de 15 ml.
- c) Los tubos se centrifugaron en un rotor de cubeta oscilante a 3000 xg durante 15 minutos.
- d) El sobrenadante se decantó cuidadosamente y se desechó sin alterar el material celular sedimentado.
- e) Las células sedimentadas se resuspendieron en 200 µl de tampón de lisis (Tris/HCl 100 mM pH 8,0, EDTA 2 mM pH 8,0, SDS al 0,5%, Triton-X-100 al 0,5%) y se mezclaron bien hasta que la disolución fue homogénea.
- 20 f) Se transfirieron 80 µl de la muestra a una placa de preparación de la muestra de 96 pocillos.
- g) Se añadieron 20 µl de proteinasa K y la disolución se incubó a 55°C durante 1 hora (este procedimiento da como resultado la lisis celular)

Extracción de ADN de muestras de orina

- 25 Se extrajo el ADN a partir de un volumen de partida de 1 ml de orina de acuerdo con QIAamp UltraSens™ Virus Handbook.

Tratamiento con bisulfito de las muestras de ADN

El tratamiento con bisulfito se realizó según el kit de modificación con bisulfito MethylEasy™ High Throughput DNA (Human Genetic Signatures, Australia) véase también más adelante.

- 30 Sorprendentemente, los autores de la presente invención han encontrado que no hay necesidad de separar el ADN microbiano de otras fuentes de ácidos nucleicos, por ejemplo cuando hay ADN microbiano en una muestra de células humanas. La etapa de tratamiento se puede utilizar para una gran mezcla de tipos diferentes de ADN y, sin embargo se puede identificar todavía mediante la presente invención un ácido nucleico específico microbiano. Se estima que los límites de detección en una mezcla compleja de ADN son los límites de la detección mediante PCR convencional que pueden deberse a una única copia de una molécula de ácido nucleico diana.

### 35 Muestras

Se puede utilizar cualquier muestra adecuada para la presente invención. Los ejemplos incluyen, pero no están limitados a, cultivos microbianos, muestras clínicas, muestras veterinarias, fluidos biológicos, muestras de cultivo de tejidos, muestras ambientales, muestras de agua, efluentes. Como la presente invención es adaptable para la detección de cualquier microorganismo, esta lista no debe considerarse exhaustiva.

### 40 Kits

La presente descripción se puede implementar en forma de diferentes kits, o combinaciones de kits e instanciar en términos de plataformas manuales, semiautomáticas o completamente robóticas. En una forma preferida, los kits MethylEasy™ o High Throughput MethylEasy™ (Human Genetic Signatures Pty Ltd., Australia) permiten la conversión de ácidos nucleicos en placas 96 o 384 utilizando una plataforma robótica tal como epMotion.

## Tratamiento con bisulfito

Un protocolo ilustrativo para el tratamiento eficaz de ácido nucleico con bisulfito se expone a continuación. El protocolo da como resultado la conservación sustancialmente de todo el ADN tratado. Este método también se denomina en la presente memoria método Human Genetic Signatures (HGS). Se apreciará que se pueden variar los volúmenes o cantidades de las muestras o de los reactivos.

El método preferido para el tratamiento con bisulfito se pueden encontrar en la Patente de los Estados Unidos US 2004219539 o el documento WO 2004096825.

A 2 µg de ADN, que pueden ser pre-digeridos con enzimas de restricción adecuadas si se desea, se les añadieron 2 µl (1/10 volumen) de NaOH 3 M (6 g en 50 ml de agua, recién elaborado), en un volumen final de 20 µl. Esta etapa desnatura las moléculas de ADN de doble hebra a una forma de cadena sencilla, ya que el reactivo de bisulfito preferiblemente reacciona con las moléculas de hebra sencilla. La mezcla se incubó a 37°C durante 15 minutos. Se puede utilizar la incubación a temperaturas por encima de la temperatura ambiente para mejorar la eficacia de la desnaturación.

Después de la incubación, se añadieron sucesivamente 208 µl de metabisulfito de sodio 2 M (7,6 g en 20 ml de agua con 416 ml de NaOH 10 N; BDH AnalaR Núm. 10356.4D; recién elaborado) y 12 µl de Quinol 10 mM (0,055 g en 50 ml de agua, BDH AnalaR Núm. 103122E; recién elaborado). El quinol es un agente reductor y ayuda a reducir la oxidación de los reactivos. También se pueden utilizar otros agentes reductores, por ejemplo, ditioneitol (DTT), mercaptoetanol, quinona (hidroquinona), u otros agentes reductores adecuados. La muestra se cubrió con 200 µl de aceite mineral. La superposición de aceite mineral evita la evaporación y la oxidación de los reactivos pero no es esencial. La muestra se incubó durante la noche a 55°C. Alternativamente, las muestras se pueden someter a ciclos en un termociclador de la siguiente manera: incubar durante aproximadamente 4 horas o durante la noche como sigue: Etapa 1, ciclación a 55°C/2 hr en la máquina de PCR; Etapa 2, 95°C/2 min. La etapa 1 se puede realizar a cualquier temperatura de aproximadamente 37°C a aproximadamente 90°C y su duración puede variar de 5 minutos a 8 horas. La etapa 2 se puede realizar a cualquier temperatura de aproximadamente 70°C a aproximadamente 99°C y su duración puede variar de aproximadamente 1 segundo a 60 minutos, o más.

Después del tratamiento con metabisulfito de sodio, el aceite se retiró, y se añadieron 1 µl de ARNt (20 mg/ml) o 2 µl de glucógeno si la concentración de ADN era baja. Estos aditivos son opcionales y se pueden utilizar para mejorar el rendimiento de ADN obtenido mediante coprecipitación con el ADN diana, especialmente cuando el ADN está presente a bajas concentraciones. Se desea generalmente el uso de aditivos como vehículo para una precipitación más eficaz de los ácidos nucleicos cuando la cantidad de ácido nucleico es < 0,5 µg.

Se realizó un tratamiento de limpieza con isopropanol como sigue: se añadieron 800 µl de agua a la muestra, se mezcló y después se añadió 1 ml de isopropanol. El agua o tampón reduce la concentración de la sal de bisulfito en el recipiente de reacción a un nivel en el que la sal no precipitará junto con el ácido nucleico diana de interés. La dilución es generalmente de aproximadamente 1/4 a 1/1000, siempre y cuando la concentración de sal se diluya por debajo de un intervalo deseado, como se describe en la presente memoria.

La muestra se mezcló de nuevo y se dejó a 4°C durante un mínimo de 5 minutos. La muestra se centrifugó en una microcentrífuga durante 10-15 minutos y el sedimento se lavó 2x con EtOH del 70%, sometiéndolo a vórtice cada vez. Este tratamiento de lavado elimina las sales residuales que precipitaron con los ácidos nucleicos.

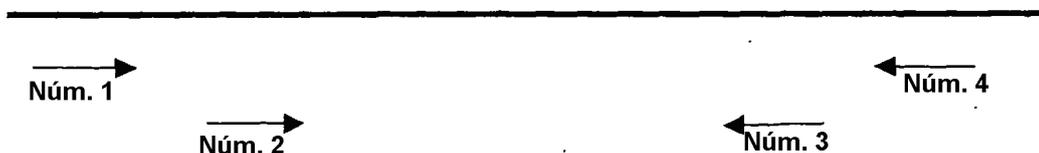
El sedimento se dejó secar y después se resuspendió en un volumen adecuado de T/E (Tris 10 mM/ EDTA 0,1 mM) pH 7,0-12,5 tal como 50 µl. Se ha encontrado que el tampón a pH 10,5 es particularmente eficaz. La muestra se incubó a 37°C a 95°C durante 1 min a 96 horas, según sea necesario para suspender los ácidos nucleicos.

Otro ejemplo de tratamiento con bisulfito se puede encontrar en el documento WO 2005021778 que proporciona métodos y materiales para la conversión de citosina en uracilo. En algunas realizaciones, un ácido nucleico, tal como ADNg, se hace reaccionar con bisulfito y un catalizador de poliamina, tal como una triamina o tetra-amina. Opcionalmente, el bisulfito comprende bisulfito de magnesio. En otras realizaciones, un ácido nucleico se hace reaccionar con bisulfito de magnesio, opcionalmente en presencia de un catalizador de poliamina y/o un catalizador de amina cuaternaria. También se proporcionan kits que se pueden utilizar para llevar a cabo los métodos de la invención. Se apreciará que estos métodos también serían adecuados para la presente invención en la etapa de tratamiento.

## Amplificación

Las amplificaciones mediante PCR se realizaron en mezclas de reacción de 25 µl que contenían 2 µl de ADN genómico tratado con bisulfito, utilizando la mezcla maestra Promega PCR, 6 ng/µl de cada uno de los cebadores. Se utilizan cebadores anidados específicos de la hebra para la amplificación. Se llevaron a cabo amplificaciones de PCR de 1ª ronda usando los cebadores de PCR 1 y 4 (véase más abajo). Después de la 1ª ronda de amplificación, se transfirió 1 µl del material amplificado a premezclas para la 2ª ronda de PCR que contenían los cebadores de

PCR 2 y 3 y se amplificaron como se ha descrito previamente. Las muestras de los productos de PCR se amplificaron en un termociclador ThermoHybaid PX2 en las condiciones: 1 ciclo de 95°C durante 4 minutos, seguido de 30 ciclos de 95°C durante 1 minuto, 50°C durante 2 minutos y 72°C durante 2 minutos; 1 ciclo de 72°C durante 10 minutos.



5

Amplificación múltiplex

Si se requiere la amplificación múltiplex para la detección, se puede llevar a cabo la siguiente metodología.

10

Se añade 1 µl de ADN tratados con bisulfito a los siguientes componentes en un volumen de reacción de 25 µl, x1 mezcla maestra Qiagen múltiplex, 5-100 ng de cada INA o cebador de oligonucleótido de 1ª ronda, MgSO<sub>4</sub> 1,5 a 4,0 mM, 400 µM de cada dNTP y 0,5-2 unidades de la mezcla de polimerasa. Los componentes se someten a ciclos en un termociclador con tapa caliente como sigue. Típicamente puede haber hasta 200 secuencias de los cebadores individuales en cada reacción de amplificación

Etapa 1	94°C	15 minutos	1 ciclo
Etapa 2	94°C	1 minuto	
	50°C	3 minutos	35 ciclos
	68°C	3 minutos	
Etapa 3	68°C	10 minutos	1 ciclo

15

A continuación se realiza una segunda ronda de amplificación sobre una alícuota de 1 µl de la primera ronda de amplificación que se transfiere a un tubo de reacción para la segunda ronda que contiene la mezcla de reacción de la enzima y cebadores de la segunda ronda apropiados. La ciclación se realiza a continuación como se ha descrito anteriormente.

Cebadores

20

Se pueden utilizar cualquier cebador de PCR adecuado para la presente invención. Un cebador tiene típicamente una secuencia complementaria a una secuencia que será amplificada. Los cebadores son típicamente oligonucleótidos, pero pueden ser análogos de oligonucleótidos.

Sondas

25

La sonda puede ser cualquier molécula de ácido nucleico o análogo de ácido nucleico adecuada. Los ejemplos incluyen, pero no están limitados a, ADN, ARN, ácido nucleico cerrado (LNA), ácido peptidonucleico (PNA), MNA, ácido altritol nucleico (ANA), ácido hexitol nucleico (HNA), ácido nucleico intercalante (INA), ácido nucleico ciclohexanilo (CNA) y mezclas de los mismos y sus híbridos, así como las modificaciones del átomo fósforo de los mismos, tales como, pero no limitados a fosforotioatos, metilfosfolatos, fosforamiditas, fosforoditiatos, fosforoselenoatos, fosfotriésteres y fosfoboranoatos. Los nucleótidos de origen no natural incluyen, pero no están limitados a los nucleótidos comprendidos en ADN, ARN, PNA, INA, HNA, MNA, ANA, LNA, CNA, CeNA, TNA, (2'-NH)-TNA, (3'-NH)-TNA, α-L-Ribo-LNA, α-L-Xilo-LNA, β-D-Xilo-LNA, α-D-Ribo-LNA, [3,2,1]-LNA, Biciclo-ADN, 6-Amino-Biciclo-ADN, 5-epi-Biciclo-ADN, α-Biciclo-ADN, Triciclo-ADN, Biciclo[4,3,0]-ADN, Biciclo[3,2,1]-ADN, Biciclo[4,3,0]amida-ADN, β-D-Ribopiranosil-NA, α-L-Lixopiranosil-NA, 2'-R-ARN, α-L-ARN o α-D-ARN, β-D-ARN. Además se pueden utilizar compuestos que no contienen fósforo para conectar nucleótidos tales como, pero no limitados a metiliminometilo, formacetato, tioformacetato y grupos conectores que comprenden amidas. En particular, los ácidos nucleicos y análogos de ácido nucleico puede comprender uno o más pseudonucleótidos intercalantes.

35

Preferiblemente, las sondas son ADN u oligonucleótidos de ADN que contienen uno o más IPN internos que forman INA.

## Electroforesis

La electroforesis de las muestras se realizó de acuerdo con la guía para el usuario del sistema E-gel ([www.invitrogen.doc](http://www.invitrogen.doc)).

## Métodos de detección

5 Existen numerosos sistemas de detección posibles para determinar el estado de la muestra deseada. Se apreciará que para la presente invención se podría utilizar cualquier sistema o método conocidos para la detección de moléculas de ácido nucleico. Los sistemas de detección incluyen, pero no están limitados a:

10 I. La hibridación de ADN marcado adecuadamente a un dispositivo de tipo micro-matriz que podría seleccionar 10 → 200.000 componentes individuales. Las matrices se podrían componer de INA, PNA o nucleótidos o matrices de de nucleótidos modificados en cualquier superficie sólida adecuada tal como vidrio, plástico, mica, nailon, cuentas, cuentas magnéticas, cuentas fluorescentes o membranas;

II. Sistemas de detección de tipo transferencia Southern;

III. Sistemas de detección mediante PCR convencional tales como gel de agarosa, lecturas fluorescentes tales como el análisis GeneScan. Análisis de hibridación sándwich, reactivos de tinción de ADN, tales como bromuro de etidio, Sybr Green, detección con anticuerpos, dispositivos del tipo lector de placa ELISA, dispositivos fluorimétricos;

15 IV. Cuantificación mediante PCR en tiempo real de fragmentos genómicos amplificados específicos o múltiples o cualquiera de sus variaciones.

V. Cualquiera de los sistemas de detección esbozados en el documento WO 2004/065625 tales como cuentas fluorescentes, productos conjugados enzimáticos, cuentas radiactivas y similares;

20 VI. Cualquier otro sistema de detección que utilice una etapa de amplificación tal como la reacción en cadena de la ligasa o tecnologías de amplificación isotérmica de ADN tales como la Amplificación de Desplazamiento de Hebra (SDA).

VII. Múltiples sistemas de detección multi-fotón.

VIII. Electroforesis y visualización en geles.

25 IX. Cualquier plataforma de detección utilizada o que se pudiera utilizar para detectar ácido nucleico.

## Ácidos nucleicos intercalantes

30 Los ácidos nucleicos intercalantes (INA) son polinucleótidos de origen no natural que pueden hibridarse a ácidos nucleicos (ADN y ARN) con especificidad de secuencia. Los INA son candidatos como alternativas/sustitutos de sondas de ácidos nucleicos en los análisis de hibridación basados en sondas debido a que exhiben varias propiedades deseables. Los INA son polímeros que se hibridan con ácidos nucleicos para formar híbridos que son termodinámicamente más estables que los correspondiente complejos ácido nucleico/ácido nucleico de origen natural. No son sustratos para enzimas que se sabe que degradan péptidos o ácidos nucleicos. Por lo tanto, los INA debe ser más estables en muestras biológicas, así como, tener una mayor vida útil que los fragmentos de ácido nucleico de origen natural. A diferencia de la hibridación de ácidos nucleicos que es muy dependiente de la fuerza iónica, la hibridación de un INA con un ácido nucleico que es bastante independiente de la fuerza iónica y se ve favorecida por la baja fuerza iónica en condiciones que desfavorecen fuertemente la hibridación de ácido nucleico natural a ácido nucleico. La fuerza de unión de INA depende del número de grupos intercalantes diseñados en la molécula, así como de las interacciones habituales de los enlaces de hidrógeno entre las bases apiladas de una manera específica en una estructura de doble hebra. La discriminación secuencia es más eficaz para el INA que reconoce ADN que para el ADN que reconoce ADN.

40 Preferiblemente, el INA es la fosforamidita de (S)-1-O-(4,4'-dimetoxitriifenilmetil)-3-O-(1-pirenilmetil)-glicerol.

Los INA se sintetizan mediante la adaptación de procedimientos de síntesis de oligonucleótidos convencionales en un formato que está disponible comercialmente. La definición completa de INA y sus síntesis se pueden encontrar en los documentos WO 03/051901, WO 03/052132, WO 03/052133 y WO 03/052134 (Unest A/S).

45 Existen en efecto muchas diferencias entre las sondas de INA y las sondas de ácido nucleico convencionales. Estas diferencias se pueden ser dividir convenientemente en diferencias biológicas, estructurales y fisicoquímicas. Como se discute anteriormente y a continuación, estas diferencias biológicas, estructurales y fisicoquímicas pueden dar lugar a resultados impredecibles cuando se intentan utilizar sondas de INA en aplicaciones en las que se han empleado tradicionalmente ácidos nucleicos. Esta no equivalencia de composiciones diferentes se observa con frecuencia en las técnicas químicas.

50

Con respecto a las diferencias biológicas, los ácidos nucleicos son materiales biológicos que desempeñan un papel central en la vida de las especies vivas como agentes de transmisión y expresión genética. Sus propiedades *in vivo* se conocen bastante bien. El INA, sin embargo, es una molécula desarrollada recientemente totalmente artificial, concebida en la mente de los químicos y elaborada utilizando la química orgánica sintética. No tiene ninguna función biológica conocida.

Estructuralmente, el INA también difiere radicalmente de los ácidos nucleicos. Aunque ambos pueden emplear nucleobases comunes (A, C, G, T y U), la composición de estas moléculas es estructuralmente diversa. Las cadenas principales de ARN, ADN e INA se componen de unidades de ribosa y 2-desoxirribosa con enlace fosfodiéster. Los INA difieren de los ADN o ARN en que tienen una o más moléculas grandes planas unidas a través de una o varias moléculas conectoras al polímero. Las moléculas planas se intercalan entre las bases en la hebra de ADN complementaria frente al INA en una estructura de doble hebra.

Las diferencias físico/químicas entre INA y ADN o ARN también son sustanciales. El INA se une al ADN complementario más rápidamente de lo que las sondas de ácido nucleico se unen a la misma secuencia diana. A diferencia de los fragmentos de ADN o ARN, el INA se une mal al ARN a menos que los grupos intercalantes estén situados en posiciones terminales. Debido a las fuertes interacciones entre los grupos intercalantes y las bases en la hebra de ADN complementario, la estabilidad del complejo INA/ADN es mayor que la de un complejo de ADN/ADN o de ARN/ADN análogo.

A diferencia de otros ácidos nucleicos tales como los fragmentos de ARN o ADN o el PNA, los INA no muestran propiedades de autoagregación o unión.

Puesto que los INA hibridan con los ácidos nucleicos con una especificidad de secuencia, los INA son candidatos útiles para desarrollar análisis basados en sondas y están particularmente adaptados para kits y análisis de escrutinio. Sin embargo las sondas de INA, no son el equivalente de las sondas de ácidos nucleicos. Por consiguiente, cualquiera de los métodos, kits o composiciones que pudieran mejorar la especificidad, sensibilidad y fiabilidad de los análisis basados en sondas serían útiles en la detección, el análisis y la cuantificación de muestras que contienen ADN. Los INA tienen las propiedades necesarias para este fin.

#### Resultados

La detección de microorganismos (tales como cepas bacterianas, virales o fúngicas) se ve a menudo dificultada por el gran número de cepas individuales de microorganismos dentro de esa especie.

Los principios *in silico* generales de la invención se ilustran usando las bacterias *Neisseria meningitidis*, *Neisseria gonorrhoeae*, *Haemophilus influenzae*, *Streptococcus sp* y *Staphylococcus* (Figuras 1 a 5). Los principios generales de la invención se han ilustrado utilizando el virus Influenza y rotavirus (Figuras 6 y 7).

Los datos bioquímicos generales para ilustrar y apoyar a la invención se describen en las Figuras 8 a 18 utilizando bacterias Gram positivas así como Gram negativas clínicamente relevantes.

#### Bacterias

La Figura 1 muestra una región de 34 nucleótidos del gen de la proteasa *iga* en *N. meningitidis* y el locus correspondiente en *N. gonorrhoeae* (ya que estas regiones existen en sus genomas bacterianos naturales) (clasificación completa; *Bacteria*; *Proteobacteria*; *Betaproteobacteria*; *Neisseriales*; *Neisseriaceae*, *Neisseria meningitidis*, Z2491 del Serogrupo A y las características completas de locus; *iga*, proteasa IgA1; GeneID 906889. Locus Tag NMA0905; Núm de acceso RefSeq. NC\_003116.1; PMID 10761919; Parkhill J et al., 2000, *Nature*, 404, 502-506). Existe una similitud de secuencia de 74% entre estas dos secuencias de 34 nucleótidos de *Neisseria*. Los cebadores basados en PCR elaborados para amplificar estas regiones en ambas especies bacterianas requerirían cebadores degenerados con 512 combinaciones posibles. La secuencia común utilizada para la parte de la amplificación mediante PCR sería la secuencia de 34 nucleótidos GYAATYW AGGYCGYCTY GAAGAYTAYA AYATGGC (SEQ ID NO: 3), donde el código convencional para designar las diferentes posiciones se proporciona más abajo: N = A, G, T o C; D = A, G o T; H = A, T o C; B = G, T o C; V = G, A o C; K = G o T; S = C o G; Y = T o C; R = A o G; M = A o C y W = A o T.

Sin embargo, cuando el ADN bacteriano de estas dos especies se trata con el reactivo de bisulfito, (que da resultado la conversión de citosinas en timinas), las secuencias de origen natural se convierten en secuencias derivadas que no tienen potencial de codificación y que no existen en la naturaleza. Las secuencias derivadas tienen ahora 97% de similitud de secuencia. Los cebadores basados en PCR diseñados para permitir la amplificación mediante PCR de estos dos loci bacterianos en un único ensayo solo requieren ahora combinaciones de solamente 2 cebadores. La combinación se basaría en la secuencia GTAATTW AGGTTGTTTT GAAGATTATA ATATGGT (SEQ ID NO: 4), donde solo la base de la posición 7 es una adenina o una timina (indicada como W). Así, la conversión con bisulfito reduce la complejidad genómica relativa de 512 a 2 tipos de cebadores. Esta reducción masiva simplifica la amplificación del mismo locus de especies bacterianas relacionadas.

Se obtienen ventajas adicionales la utilización opcional de sondas de INA para amplificar regiones de estas dos especies bacterianas, utilizando de nuevo el mismo locus. La Figura 2 ilustra la misma región de 34 nucleótidos de los genes *iga* de *N. meningitidis* y *N. gonorrhoeae* como se representa en la Figura 1, con la demostración adicional del grado en que se pueden reducir aún más la longitud y la complejidad de la sonda utilizando sondas de INA. Una secuencia de 16 unidades de INA corta AGGYCGYCTY GAAGAY (SEQ ID NO: 5) requeriría 16 posibles combinaciones de cebadores para detectar esta región, pero después de la conversión con bisulfito, sería suficiente una única secuencia cebadora, AGGTTGTTTT GAAGAT (SEQ ID NO: 6). La ventaja de la molécula de INA es que, debido a los pseudonucleótidos intercalantes que se incorporan en su cadena principal, la hibridación con el locus correcto se distingue mucho más fácilmente de la unión no específica, debido al incremento de la Tm del INA con respecto al oligonucleótido convencional. Se apreciará, sin embargo, que los oligonucleótidos convencionales todavía funcionarán adecuadamente.

Cuando las especies bacterianas estrechamente relacionadas causan síntomas clínicos similares, ADN convertido con bisulfito se puede utilizar de nuevo para diseñar sondas más simples para analizar la presencia de tipos bacterianos específicos. La Figura 3 muestra los alineamientos de ADN del gen *iga* en tres especies bacterianas, una de las cuales, *Haemophilus influenzae* es de un grupo taxonómico diferente. El tratamiento con bisulfito del ADN bacteriano dio como resultado un número mucho más pequeño de combinaciones de sondas. Esta comparación ilustra la importancia de poder analizar especies no relacionadas en un ensayo. Tanto *N. meningitidis* como *H. influenzae* causan meningitis, por lo que es ventajoso poder analizar en un ensayo todos los microbios que causan los mismos síntomas clínicos.

El análisis de un gran número de diferentes especies bacterianas del mismo grupo taxonómico se facilita de nuevo por medio de la presente invención. La Figura 4 muestra un segmento de 40 nucleótidos del gen *tuf* en 10 especies de bacterias del grupo *Streptococcus* a saber *S. oralis*, *S. mitis*, *S. dysgalactiae*, *S. cristatus*, *S. gordonii*, *S. parauberis*, *S. pneumoniae*, *S. bovis*, *S. vestibularis* y *S. uberis*. Esta región tiene aproximadamente 68% de similitud de secuencia entre las 10 especies y requiere 12.288 combinaciones de cebadores para analizar simultáneamente las 10 especies en el ensayo. La secuencia convertida con entre estas especies tiene 85% de similitud de secuencia y ahora sólo requiere 64 posibles combinaciones de cebadores.

El análisis de diferentes cepas que pertenecen a la misma especie bacteriana se simplifica también mediante la invención. La Figura 5 ilustra un segmento de 23 nucleótidos del gen de enterotoxina *se de Staphylococcus aureus*. La secuencia natural de esta región del gen tiene sólo 56% de similitud de secuencia entre las 7 cepas y requiere 1536 combinaciones de cebadores, mientras que la secuencia convertida con bisulfito tiene 74% de similitud de secuencia y sólo requiere de 64 combinaciones de cebadores.

#### Análisis de ácidos nucleicos virales y reducción de la complejidad genómica relativa

El principio de reducción de la complejidad genómica relativa también se puede aplicar a los grupos virales, tales como el virus Influenza que tiene un genoma de ADN, así como a grupos virales que tienen genomas de ARN, (ya que el ARN se puede convertir en ADN por medio de la transcriptasa inversa y después el consiguiente tratamiento con bisulfito). Para ilustrar la aplicación para la detección viral, se han utilizado el gen de la neuraminidasa de las cepas de virus Influenza, (familia Orthomyxoviridae), y la proteína de superficie que codifica el gen de VP4 de las cepas de rotavirus, (familia Reoviridae), teniendo ambos virus un genoma de ARN segmentado. La taxonomía de los virus de la influenza es compleja, estando basados por ejemplo los tipos A, B y C, en características antigénicas, y estando basados los subtipos adicionales en el lugar de origen, el año de aislamiento, el número de producto aislado y el subtipo. Esto refuerza la necesidad en primera instancia de poder identificar los virus de influenza como grupo, y sólo entonces profundizar para analizar los sub-subniveles de clasificación.

La taxonomía de los rotavirus también es compleja. El número de serotipos de rotavirus es grande, siendo reconocidos dos serotipos principales, los serotipos P y G. Hay mínimamente 14 diferentes serotipos G y su detección inequívoca tiene importancia en la medicina pediátrica. Se estima que a la edad de tres años, casi todos los niños en todo el mundo ya se ha infectado por lo menos una vez por Rotavirus, a pesar de que estas infecciones pueden ser subclínicas y sólo tienen leves efectos sobre el tracto gastrointestinal.

Las consecuencias de la infección por influenza en el nivel clínico son bien conocidas, con una significativa morbilidad y mortalidad casi cada invierno. Sin embargo puede haber complicaciones secundarias masivas después de la infección, especialmente por *Streptococcus pneumoniae*, *Haemophilus influenzae* y *Staphylococcus aureus*. Es muy claramente ventajoso ser capaz de analizar simultáneamente para ambas infecciones virales e infecciones bacterianas ya que pueden surgir complicaciones neumónicas de características mixtas de infecciones bacterianas y virales, y el inmediato tratamiento con antibióticos puede ser una terapia eficaz.

La reducción de la complejidad genómica relativa en 9 cepas de influenza diferentes se muestra en la Figura 6. Una región de 20 nucleótidos del gen de la neuraminidasa del virus de la influenza se muestra en su forma de ADN. Hay 50% de similitud de secuencia entre estos 9 productos aislados.

Después de la conversión con bisulfito, la similitud de secuencia se ha incrementado a 75%. En su forma original, que requerirían 2.048 posibles combinaciones de cebadores para analizar estas 9 cepas, mientras que después de la conversión con bisulfito solamente se necesitan 48 combinaciones de cebadores.

5 La reducción complejidad genómica relativa del gen de VP4 de 3 diferentes cepas de rotavirus se muestra en la Figura 7. Una región de 20 nucleótidos del gen de VP4 tiene 52% de similitud de secuencia antes de la conversión y el 74% después de la conversión. El número de combinaciones de cebadores se reduce de 512 a 32.

Los datos moleculares que apoyan el enfoque in silico de la simplificación de los genomas microbianos, como medio de detección de microorganismos se ilustra en las Figuras 8 a 15 utilizando especies microbianas clínicamente relevantes que se encuentran comúnmente en hospitales y unidades de patología de prueba.

10 Es una clara ventaja, y un imperativo clínico para la detección rápida de microorganismos contaminantes, que se pudiera tomar una decisión inicial entre la presencia de bacterias Gram positivas o Gram negativas en una muestra. El método descrito en la presente memoria proporciona dicho ensayo utilizando los genes ribosomales 23S de diferentes especies de bacterias para generar un conjunto de cebadores que permiten detectar las bacterias Gram  
15 positivas o Gram negativas mediante la utilización de tales cebadores sobre genomas simplificados a través de una reacción de amplificación. Las secuencias de 23S son ideales para tales distinciones de alto nivel, puesto que se encuentran en todas las especies bacterianas, a diferencia de algunas secuencias codificantes de proteínas que son adiciones opcionales a algunos genomas bacterianos, tal como se observa en el ejemplo de *S. galactiae* anterior. Muchas secuencias microbianas codificantes de proteínas son similares a "restos y desechos" genómicos, y su  
20 utilidad radica en la diferenciación entre las categorías taxonómicas de nivel inferior, tales como diferentes cepas microbianas, tipos o productos aislados, o en el caso de los virus, entre diferentes tipos o mutaciones recién surgidas. Las secuencias genómicas normales y simplificadas de estos dos componentes, los genes de ARN ribosomal que no codifican proteínas, y el gen recA que codifica proteína de bacterias se proporcionan en las Figuras 15 y 16, respectivamente. Las secuencias de los cebadores utilizados para realizar las reacciones de  
25 amplificación para los amplicones bacterianos 23S se proporcionan en la Tabla 1. Las secuencias de los cebadores utilizadas para realizar las reacciones de amplificación para los amplicones recA se proporcionan en la Tabla 2. Todos los cebadores se elaboran para el ADN tratado con bisulfito y se muestran en la orientación 5' a 3'.

La Tabla 1 presenta las secuencias adecuadas de cebadores bacterianos usadas para amplificar el ADN simplificado con bisulfito del gen o los genes de ARN ribosomal 23S utilizando alineamientos para generar  
30 cebadores para la detección Gram positivos (Pos), Gram negativos (Neg). Además se diseñaron los cebadores para la detección específica de *Mycoplasma spp* (Myc), *Staphylococcus spp* (Staph), *Streptococcus spp* (Strep), *Neisseria spp* (NG), *Chlamydia* (CT), y *Escherichia coli* y *Klebsiella pneumoniae* (EC).

Los siguientes símbolos designan las adiciones de bases siguientes: N = A, G, T o C; D = A, G o T; H = A, T o C; B = G, T o C; V = G, A o C; K = G o T; S = C o G; Y = T o C; R = A o G; M = A o C y W = A o T.

Todos los cebadores utilizados se basaron en secuencias de ADN simplificadas con bisulfito.

35 **Tabla 1** Cebadores bacterianos

Cebadores 23S	Secuencia 5'-3'	SEQ ID NO
Pos-R1F1	GGTTTTTTTTGAAATAGTTTTAGGGTTA	7
Neg-R1F1	GGTTTTTTTTGAAARTTATTTAGGTAGT	8
Pos-R1F2	TGGKAGTTAGAWTGTGRRWGATAAG	9
Neg-R1 F2	TGGGAGATAKATRGTGGGTGTTAAT	10
Pos-R1F3	GGATGTGGDRTTKTKWAGATAA	11
Neg-R1 F3	TGAWGTGGGAAGGTWTAGATAG	12
Pos-R1R1	HCAATMHHACTTCAMMMCMYT	13
Neg-R1R1	WCAAHHCACCTTCAHMAACYTAC	14

ES 2 399 054 T3

Cebadores 23S	Secuencia 5'-3'	SEQ ID NO
Pos-R1R2	ACCAACATTCTCACTYMTAAWMAMTCCAC	15
Neg-R1R2	ATCAACATTCACACTTCTAATACCTCCAA	16
W-Pos-R1F1	GGTTTTTTTTYGAAATAGTTTTAGGGTTA	17
W-Neg-R1F1	GGTTTTTTTTYGAAARTTATTTAGGTAGT	18
W-Pos-R1F2	YGGKAGTTAGAWYGYGRRWGATAAG	19
W-Neg-R1F2	YGGGAGATAKAYRGYGGGTGTTAAT	20
W-Pos-R1F3	GGATGTGGDRTTKYKWAGATAA	21
W-Neg-R1F3	YGAWGTGGGAAGGTWTAGATAG	22
W-Pos-R1R1	HCRATMHHRCTTCRMMMCMMYT	23
W-Neg-R1R1	WCRAHHCACCTTCAHMRACYTAC	24
W-Pos-R1R2	ACCRACATTCTCACTYMTAAWMAMTCCAC	25
W-Neg-R1R2	ATCAACATTCRCACCTTCTAATACCTCCAA	26
Pos-R2F1	KTRAGAAAAGTWTTTAGDDAGRK	27
Neg-R2F1	TTTARGAAAAGTTWTTAAGTWTTA	28
Pos-R2F2	AGDTRAGRWGAGDATTTTWAGGTKR	29
Neg-R2F2	GGKTRGGWWGAGAATWTTAAGGTGT	30
Pos-R2R1	AATYTMYMATTAACAATAACMCAA	31
Neg-R2R1	AATCTCAAWAAAAACAAYMYMACC	32
Pos-R2R2	ACMHACATCTTACWMAYAYTAYAAYTTCACC	33
Neg-R2R2	MAYTACATCTTACAACMAHWTCAYTTCACT	34
Pos-R2R3	CMATAYYAAAYTACAATAAACTC	35
Neg-R2R3	CAATAYMAAACTAYAATAAAAATT	36
Pos-R3F1	GGTGAARTTRTARTRTKWGTGAAGATGTDKG	37
Neg-R3F1	AGTGAARTTGAWDTKGTTGTGAAGATGTART	38
Pos-R3F2	GATWGGATGGAAAGATTTTRTRGAG	39

ES 2 399 054 T3

Cebadores 23S	Secuencia 5'-3'	SEQ ID NO
Neg-R3F2	KGTWAGATGGAAAGATTTTGTGAAT	40
Pos-R3R1	HYMAYMMWAYHAAAATAATATCC	41
Neg-R3R1	TCAAMMMYWMMAAAATAATATTT	42
Pos-R3R2	AWCCATTCTAAAAAACCTTTAAACA	43
Neg-R3R2	AACCAWWMYWAAMHMACCTTCAWACT	44
CE-F1	GTTGGTAAGGTGATATGAATTGTTATAA	45
CE-F2	TTATTATTAATTGAATTTATAGGTTA	46
CE-F3	GAGGAGTTTAGAGTTTGAATTAGTRTG	47
EC_R1	TATATACAAAACCTATCACCCCTATATC	48
CE-R2	TCATCAAACCTCACAACAYATAC	49
NG-F1	TTGAGTAAGATATTGATGGGGGTAA	50
NG-F2	TATGGTTAGGGGGTTATTGTA	51
NG-R1	AATCTATCATTTAAAACCTTAACC	52
NG-R2	CCTAACTATCTATACCTTCCCACT	53
NG-R3	CACTCCCCTACCATACCAATAAACC	54
CT-R1F1	GTATGATGAGTTAGGGAGTTAAGTTAAA	55
CT-R1F2	GGTGAGGTTAAGGGATATATA	56
CT-R1F3	AAAAGAGTGAAGAGTTGTTTGGTTTAGATA	57
CT-R1R1	TCCAAACCTTTTTCAACATTAAC	58
CT-R1R2	CCCTAAAATTATTTCAAAAAAACA AAA	59
CT-R2F1	TTAGTGGGGGTTTATTGGTTTATTAATGGA	60
CT-R2F2	TAAGGAAGTGATGATTTGAAGATAGTTGGA	61
CT-R2R1	ACACCTTCTCTACTAAATACT	62
CT-R2R2	TATACCATAAATCTTCACTAATATC	63
CT-R3F1	TTGTGTAGATGATGGAGTAGTAGGTTA	64

ES 2 399 054 T3

Cebadores 23S	Secuencia 5'-3'	SEQ ID NO
CT-R3F2	GAATGATGGAGTAAGTTAAGTATGTGGA	65
CT-R3R1	TAAAAATTATTTCTTAAAAACCTCACT	66
CT-R3R2	AAATTATCTCACACACCTTAAAATAT	67
CT-R4F1	AATGTTAAAAGGTTAAAGGGATAT	68
CT-R4F2	TATTGAATTTAAGTTTTGGTGAATGGTT	69
CT-R4R1	CCAATATTTCAACATTAACCTCCCACTCTC	70
CT-R4R2	ATATCCATCTTCCAAATTCATAAAATAAT	71
CT-R4R3	TAAACAACAACAATTCCACTTTCC	72
Myc-R1F1	ATAGGAAAAGAAAWTGAAWGWGATTTTG	73
Myc-R1 F2	GTGTAGTGGTGAGTGAAAGTGGAATAGG	74
Myc-R1R1	TAAACAAMTTCMMTCAAATAACATTTYCAA	75
Myc-R1R2	CTAATTAATATTTAAACTTACCC	76
Myc-R2F1	TTTTGAAATTATATGTTTATAATGT	77
Myc-R2F2	AAGTATGAGTTGGTGAGTTATGATAGT	78
Myc-R2R1	CCTCCAMTTAWTYATAATCTYAC	79
Myc-R2R2	CACCWAAAAYACACCATCATAATT	80
Myc-R3F1	TGTAGTTAGATAGTGGGGTATAAGTTTTA	81
Myc-R3F2	AGGGGAAGAGTTTAGATTATTA	82
Myc-R3R1	ATAACTTCAWCYCMWATAACAACACTCAT	83
Myc-R3R2	ATCAATTTAAAAAATTCTCACTCYCAA	84
Myc-R4F1	TTTTTATWATTGGATTTGGGGWTAAA	85
Myc-R4F2	TKKTWWTAGTATTGAGAATGA	86
Myc-R4F3	TGTAAATTWATTTTGTAAAGTTWGT	87
Myc-R4F4	GAATGAGGGGGGATTGTTTAATT	88
Myc-R4R1	TCTATAACCAAAAACAATCAAAAAATA	89

ES 2 399 054 T3

Cebadores 23S	Secuencia 5'-3'	SEQ ID NO
Myc-R4R2	CATTACACCTAACAAATATCTTCACC	90
Myc-R5F1	ATWWATAGGTTGAATAGGTRAGAAAT	91
Myc-R5F2	ATAGTGATTTGGTGGTTTAGTATGGAAT	92
Myc-R5R1	CAAACCTACTTCAACTCAAAAATAAAAATAAAT	93
Myc-R5R2	ACAACAATTTAAACCCAACTCACATATCT	94
Myc-R5R3	AAAAYAAMWCTYTTCAATCTTCCTAYAAA	95
Strep-R1F1	ATWWTGTTAAGGDWRTGARRAGGAAG	96
Strep-R1F2	TAGRAGGGTAAATTGARGWGTTTA	97
Strep-R1F3	TKATTTGGGAARRTWRGTTAAAGAGA	98
Strep-R1 R1	TCTCTCAACTTAACCTCACATCAT	99
Strep-R1R2	ATAATTTCAAATCTACAWCMWAAT	100
Strep-R2F1	RATKTATTGGAGGATTGAATTAGGG	101
Strep-R2F2	ATGTTGAAAAGTGTTTGGATGAT	102
Strep-R2R1	TCTAAAATYAATAAWCCAAAATAAMCCCCTC	103
Strep-R2R2	ACTACCAAYHATAWHTCATTAAC	104
Strep-R3F1	AGGTTGAKATTTTTGTATTAGAGTA	105
Strep-R3F2	RWAGTGATGGAGGGATGTAGTAGGTTAAT	106
Strep-R3R1	CTTTTCTYAACAATATAACATCACT	107
Strep-R3R2	CTCTCAMTCACCTAAAACACTCA	108
Staph-R1F1	AGAAGTTGATGAAGGATGTTATTAATGA	109
Staph-R1 F2	GTTATTGATATGTGAATWTATAGTATRTT	110
Staph-R1R1	CAAAAYHTTACCTTCTYTAATYC	111
Staph-R1 R2	CAACAAAATTYCACATACTCCAT	112
Staph-R2F1	GATTTGATGTAAGGTTAAGTAGT	113
Staph-R2F2	TTGGTTAGGTTGAAGTTTAGGTAATATTGAA	114

ES 2 399 054 T3

Cebadores 23S	Secuencia 5'-3'	SEQ ID NO
Staph-R2F3	GATTTATGTTGAAAAGTGAGTGGATGAATTGA	115
Staph-R2R1	CCTYTTTCTAACTCCCAAATTAATTAAT	116
Staph-R3F1	GAAGTTGTGGATTGTTTTTTGGATA	117
Staph-R3F2	AAGGGTGTGAAAGTATGATTGTAAGGATAT	118
Staph-R3R1	TACAMTCCAAYMACACACTTCACCTATCCTA	119
Staph-R3R2	CAACAATATAAAATCAACAACACTCAA	120
Staph-R4F1	AGGAGTGGTTAGTTTTTTGTGAAGTTA	121
Staph-R4F1	ACAAATTAATAAAWCCAACACAACACT	122
Staph-R4F2	TAACACTATCTCCCACCAYAATMAAT	123

5 La Tabla 2 muestra las secuencias de cebadores bacterianos utilizadas para amplificar el ADN simplificado de del gen que codifica la proteína recA utilizando alineamientos de *Staphylococcus aureus* (SA), *Staphylococcus epidermidis* (SE), *Serratia marscesens* (SM), *Escherichia coli* (CE) y *Yersinia enterocolitica* (YE) para la tipificación bacteriana única.

Tabla 2 Secuencias de cebadores bacterianos utilizadas en la amplificación de ADN simplificado del gen que codifica la proteína recA

Específico de recA	Secuencia	SEQ ID NO
A-SA-F1	TAGGTTGTTGAGTTTTAATTATA	124
A-SA-F2	GAAGTATAAAGTAATGGTGGGGTG	125
A-SA-R1	TACAATATCAACTACACCACTTCTAACAAT	126
A-SA-R2	TAATAAAAATAACAATTATATTT	127
A-SE-F1	AAGGTTGTAGAGTATTAAGTATTTAAG	128
A-SE-F2	GTTGATAATGTATTAGGGGTTGGA	129
A-SE-F3	ATATGGATTTGAAAGTTTAGGTAAGATG	130
A-SE-R1	TACTACTAAATCAACAACAACAATATCCACA	131
A-SE-R2	CTTAATACTTAAAACATTAATCT	132
A-SM-F1	GAGAATAAGTAAAAGGTGTTAGTTGTG	133
A-SM-F2	GATTTTTATTGGTTTATTGTTATTTGATATTGTT	134

Específico de recA	Secuencia	SEQ ID NO
A-SM-R1	CAAATAATCAATATCAACACCCCACTTTTTTC	135
A-SM-R2	TACACACCACCAAACCCATATAC	136
A-CE-F1	GAAAATAAATAGAAAAGTGTGGTG	137
A-CE-F2	TGTTTTTATTGGATATTGTGTTT	138
A-CE-R1	CAATAACATCTACTACACCAAAAACAC	139
A-CE-R2	CATATTAACACTTCAAATTACCC	140
A-YE-F1	TATGTGTTTTGGTGAAGATTGTTA	141
A-YE-F2	TTTTGATATTGTATTGGGGGTG	142
A-YE-F3	GGTTTGTTAATGGGGTGTATTGTTGAG	143
A-YE-R1	CATACTCTACATCAATAAAA	144

La Tabla 1 muestra las secuencias de cebadores bacterianos utilizadas para amplificar ADN simplificado con bisulfito del gen o los genes de ARN ribosomal 23S utilizando alineamientos múltiples para generar cebadores óptimos para la detección de bacterias Gram positivas (indicadas como Pos), y Gram negativas (indicadas como Neg). Además también se diseñaron los cebadores para la detección específica de los grupos de especies, así como para las especies individuales. Las denominaciones para estos grupos de cebadores bacterianos son las siguientes; *Escherichia coli* y *Klebsiella pneumoniae* (EC), *Neisseria* spp (NG), *Chlamidia* (CT), *Mycoplasma* spp (Myc), *Streptococcus* spp (Strep) y *Staphylococcus* spp (Staph). Las subdenominaciones F y R hacen referencia a los cebadores directos e inversos, respectivamente. Además, cuando es necesaria más de una posible base en una posición determinada de nucleótido, la degeneración de base está dada por el código siguiente: N = A, G, T o C; D = A, G o T; H = A, T o C; B = G, T o C; V = G, A o C; K = G o T; S = C o G; Y = T o C; R = A o G; M = A o C; y W = A o T. Para reiterar, todos los cebadores usados en esta invención se basan en secuencias de ADN simplificadas con bisulfito.

La Tabla 2 muestra las secuencias de cebadores bacterianos utilizadas para amplificar ADN modificado con bisulfito del gen que codifica la proteína recA utilizando alineamientos de *Staphylococcus aureus* (SA), *Staphylococcus epidermidis* (SE), *Serratia marscesens* (SM), *Escherichia coli* (CE) y *Yersinia enterocolitica* (YE) para una tipificación bacteriana única.

La Figura 8 muestra los productos de amplificación obtenidos mediante PCR a partir de las regiones del gen ribosomal 23S genómicamente simplificadas de bacterias Gram positivas y Gram negativas, detectándose los amplicones de tamaño apropiado como bandas de longitud específica mediante electroforesis en gel de agarosa. La flecha indica el tamaño esperado de los amplicones con respecto a los marcadores de tamaño patrón que se ejecutan en la calle del Marcador, (M). El uso de cebadores específicos para las bacterias Gram negativas revela bandas sólo en las seis calles Gram negativas 1 a 6, (panel superior), para *Escherichia coli*, *Neisseria gonorrhoeae*, *Klebsiella pneumoniae*, *Moraxella catarrhalis*, *Pseudomonas aeruginosa* y *Proteus vulgaris*. El uso de cebadores específicos para las bacterias Gram positivas revela sólo las bandas en las seis calles Gram positivas, 7 a 12 (panel inferior) para *Enterococcus faecalis*, *Staphylococcus epidermidis*, *Staphylococcus aureus*, *Staphylococcus xylois*, *Streptococcus pneumoniae* y *Streptococcus haemolyticus*.

La Figura 9 muestra los productos de amplificación obtenidos mediante PCR a partir de las regiones del gen ribosomal 23S genómicamente simplificadas diseñadas para detectar amplicones de sólo dos especies de bacterias Gram negativas, (en este ejemplo) *E. coli* y *K. pneumoniae*. La especificidad de la metodología de amplificación se ilustra por la presencia de amplicones en las calles 1 y 3, que representan *E. coli* y *K. pneumoniae*, y la ausencia de productos de amplificación en el calle 2, así como de las calles 4 a 12, representado estas 10 calles vacías las restantes 10 especies de bacterias utilizadas en el ensayo.

La Figura 10 muestra los productos de amplificación obtenidos mediante PCR a partir de las regiones del gen ribosomal 23S genómicamente simplificadas usando cebadores específicos para un solo grupo bacteriano, *Neisseria*. La especificidad de la metodología de simplificación genómica se ilustra por la presencia de un amplicón solo en el calle 2, que representa *Neisseria gonorrhoeae*, y la ausencia de un producto de amplificación en el calle 1, así como de las calles 3 a 12, representando estas 11 calles vacías las 11 especies de bacterias restantes utilizadas en el ensayo.

Para el análisis de las especies microbianas individuales, también se pueden utilizar cuando sea apropiado genes que codifican proteínas, con la condición de que las diferentes cepas de microorganismos no sean polimórficas para la presencia/ausencia de la secuencia del gen en cuestión.

La Figura 11 ilustra el uso de cebadores para el gen *recA* bacteriano de *E. coli*. La especificidad de la amplificación se ilustra por la presencia del amplicón de tamaño correcto en la calle 1 y su ausencia de las calles restantes 2 a 12, lo que representa otras 11 especies de bacterias.

Los datos de la Figura 12 ilustran adicionalmente la especificidad de los cebadores que revela la pertenencia a un grupo más grande de bacterias, tales como estafilococos. Los productos de amplificación obtenidos mediante PCR de las regiones del gen ribosomal 23S genómicamente simplificadas usando cebadores específicos para los estafilococos revelan amplicones solo en las calles 8, 9, y 10, representando *Staphylococcus epidermidis*, *Staphylococcus aureus* y *Staphylococcus xylois*. La ausencia de un producto de amplificación en las calles 1 a 7, así como de las calles 11 y 12, demuestran la especificidad de la reacción. Las 9 calles vacías representan las 9 especies de bacterias estafilocócicas no utilizadas en el ensayo.

La Figura 13 muestra los productos de amplificación obtenidos mediante PCR a partir de las regiones del gen ribosomal 23S genómicamente simplificadas usando cebadores específicos para las bacterias Estreptocócicas.

Los productos de amplificación obtenidos mediante PCR de las regiones del gen ribosomal 23S genómicamente simplificadas usando cebadores específicos para Streptococci revela amplicones sólo en las calles 11 y 12, representando *Streptococcus pneumoniae* y *Streptococcus haemolyticus*. La ausencia de un producto de amplificación en las calles 1 a 10, revela la especificidad de la reacción. Estas 10 calles vacías representan las 10 especies de bacterias no Estreptocócicas utilizadas en el ensayo.

La Figura 14 muestra los productos de amplificación obtenidos mediante PCR de un gen que codifica una proteína de la región genómicamente simplificada del gen *recA* de *Staphylococcus epidermidis*, (Calle 8). Las dos bandas (flechas) representan los amplicones de transporte de las amplificaciones de PCR de la primera ronda (banda superior) y la segunda ronda (banda inferior). La ausencia de amplicones en las calles 1 a 7 y 9 a 12 muestra la especificidad del método y destaca el punto en el que se pueden utilizar los genes que codifican proteínas en circunstancias particulares en lugar de los componentes no codificantes del genoma, para lograr la detección de una sola especie de bacterias.

La Figura 15 muestra la detección de amplicones usando cebadores específicos dirigidos a los genes ribosomal 23S genómicamente simplificados de Chlamydia. Las reacciones de PCR se llevaron a cabo por duplicado debido a la baja cantidad de ADN de partida. La calle número 5 era ADN extraído de la orina de un individuo negativo conocido. La presencia de una banda en cualquiera de los duplicados se consideró una reacción positiva de la presencia de ADN de Chlamydia.

La Figura 16 muestra la secuencia de nucleótidos normal del gen de ARN ribosomal 23S de *E. coli* y la misma secuencia después de la simplificación genómica, donde para fines ilustrativos todas las citosinas se han sustituido por timinas.

La Figura 17 muestra la secuencia de nucleótidos normal del gen *recA* de *E. coli* y la misma secuencia después de la simplificación genómica, donde para fines ilustrativos todas las citosinas se han sustituido por timinas.

En resumen, el ADN tratado con bisulfito de las fuentes microbianas, cuando se amplifica usando cebadores genómicamente simplificados, ya sean oligonucleótidos o ácidos nucleicos modificados tales como INA proporcionan un sistema de detección sin igual para la búsqueda de microorganismos de cualquier tipo dentro de una muestra, ya sea la muestra de material clínico humano o en el otro extremo de una fuente ambiental tal como agua contaminada. La presente invención se ha demostrado para una amplia gama de diferentes especies de bacterias, y por un virus clínicamente relevante. La detección de microorganismos eucarióticos unicelulares tales como la levadura *Saccharomyces cerevisiae* o sus parientes es una simple ampliación del método. Ésta requiere fuentes de secuencias genómicas similares, tales como las secuencias ribosomales 18 o 28S, o como se ha mostrado, secuencias codificantes de proteínas que son específicas para una especie, tipo, cepa o mutante o polimorfismo dados.

Las implicaciones prácticas del sistema de detección de acuerdo con la presente invención también son importantes. Mientras que los principios descritos en detalle en la presente memoria han demostrado usando PCR para la

5      amplificación, las lecturas se pueden desarrollar a través de cualquier método conocido en la técnica. Con el énfasis actual en los sistemas de detección de micromatrices, se podría detectar una mayor variedad de microorganismos con ADN genómicamente simplificado ya que el tratamiento con bisulfito reduce la complejidad genómica y por lo tanto permite someter a ensayo más clases de microorganismos en las micromatrices con un menor número de detectores (características).

10     Si por ejemplo se fuera a construir una matriz para detectar aproximadamente 250.000 microorganismos diferentes en un ensayo, la metodología actual no podría proporcionar una plataforma de detección pragmática adecuada, ya que vería abrumada por las limitaciones físicas de la plataforma detectora. Sin embargo, con la simplificación genómica, una pequeña micromatriz podría detectar más o menos 1000 diferentes categorías bacterianas de alto nivel. Los positivos de dicho ensayo se podrían evaluar a continuación utilizando otra matriz, que contuviera simplemente los representantes de los grupos que dieron positivo en el ensayo inicial.

LISTA DE SECUENCIAS

5	<110> Human Genetic Signatures	
	<120> Detección de Microorganismos	
	<130> 205591675	
10	<160> 226	
	<170> PatentIn versión 3.1	
	<210> 1	
15	<211> 14	
	<212> DNA	
	<213> artificial	
	<400> 1	
20	atatatat at	14
	<210> 2	
	<211> 14	
	<212> DNA	
25	<213> artificial	
	<400> 2	
	aaaaaattt tttt	14
30	<210> 3	
	<211> 34	
	<212> DNA	
	<213> Neisseria	
35	<400> 3	
	gyaatywagg ycgycytaa gaytayaaya tggc	34
	<210> 4	
	<211> 34	
40	<212> DNA	
	<213> artificial	
	<400> 4	
45	gtaattwagg ttgtttgaa gattataata tgg	34
	<210> 5	
	<211> 16	
	<212> DNA	
	<213> artificial	
50	<400> 5	
	aggycgycty gaagay	16
	<210> 6	
55	<211> 16	
	<212> DNA	
	<213> artificial ,	
	<400> 6	
60	aggttgttt gaagat	16
	<210> 7	
	<211> 28	

ES 2 399 054 T3

	<212> DNA <213> bacteriano	
5	<400> 7 ggttttttt gaaatagttt tagggta	28
10	<210> 8 <211> 28 <212> DNA <213> bacteriano	
15	<400> 8 ggttttttt gaaartatt taggtagt	28
20	<210> 9 <211> 25 <212> DNA <213> bacteriano	
25	<400> 9 tggkagttag awtgrrrwg ataag	25
30	<210> 10 <211> 25 <212> DNA <213> bacteriano	
35	<400> 10 tgggagatak atrgtgggtg ttaat	25
40	<210> 11 <211> 22 <212> DNA <213> bacteriano	
45	<400> 11 ggatgtggdr tktkwagat aa	22
50	<210> 12 <211> 22 <212> DNA <213> bacteriano	
55	<400> 12 tgawgtggga aggtwtagat ag	22
60	<210> 13 <211> 22 <212> DNA <213> bacteriano	
65	<400> 13 hcaatmhhac ttcammcm ym yt	22
70	<210> 14 <211> 23 <212> DNA <213> bacteriano	
75	<400> 14 wcaahhcacc tcahmaacy tac	23
	<210> 15 <211> 29	

ES 2 399 054 T3

	<212> DNA	
	<213> bacteriano	
	<400> 15	
5	accaacattc tcactymtaa wmamtccac	29
	<210> 16	
	<211> 29	
	<212> DNA	
10	<213> bacteriano	
	<400> 16	
	atcaacattc acactctaa tacctcaa	29
15	<210> 17	
	<211> 28	
	<212> DNA	
	<213> bacteriano	
20	<400> 17	
	ggtttttty gaaatagttt tagggta	28
	<210> 18	
	<211> 28	
25	<212> DNA	
	<213> bacteriano	
	<400> 18	
30	ggtttttty gaaartatt taggtagt	28
	<210> 19	
	<211> 25	
	<212> DNA	
	<213> bacteriano	
35	<400> 19	
	yggkagttag awygyrrwg ataag	25
40	<210> 20	
	<211> 25	
	<212> DNA	
	<213> bacteriano	
45	<400> 20	
	ygggagatak ayrgygggtg ttaat	25
	<210> 21	
	<211> 22	
	<212> DNA	
50	<213> bacteriano	
	<400> 21	
	ggatgtgdr tkykwagat aa	22
55	<210> 22	
	<211> 22	
	<212> DNA	
	<213> bacteriano	
60	<400> 22	
	ygawgtggga aggtwtgat ag	22
	<210> 23	
	<211> 22	

	<212> DNA <213> bacteriano	
5	<400> 23 hcratmhrc tcrmmcm y	22
10	<210> 24 <211> 23 <212> DNA <213> bacteriano	
15	<400> 24 wcrAhcacc tcahmrcy tac	23
20	<210> 25 <211> 29 <212> DNA <213> bacteriano	
25	<400> 25 accracattc tcactymta wamtccac	29
30	<210> 26 <211> 29 <212> DNA <213> bacteriano	
35	<400> 26 atcaacattc racttctaa tacctcaa	29
40	<210> 27 <211> 24 <212> DNA <213> bacteriano	
45	<400> 27 kttragaaaa gtwttagdd agrk	24
50	<210> 28 <211> 24 <212> DNA <213> bacteriano	
55	<400> 28 ttargaaaa gttwtaagt wta	24
60	<210> 29 <211> 25 <212> DNA <213> bacteriano	
	<400> 29 agdtrgrwg agdatttwa ggtr	25
60	<210> 30 <211> 25 <212> DNA <213> bacteriano	
	<400> 30 gktrggwwg agaawttaa ggtg	25
	<210> 31 <211> 25	

	<212> DNA <213> bacteriano	
5	<400> 31 aatytmymat taaaacaata cmcaa	25
10	<210> 32 <211> 25 <212> DNA <213> bacteriano	
15	<400> 32 aatctcaaaw aaaaacaaym ymacc	25
20	<210> 33 <211> 32 <212> DNA <213> bacteriano	
25	<400> 33 acmhacatct tcacwmayay tayaaytca cc	32
30	<210> 34 <211> 32 <212> DNA <213> bacteriano	
35	<400> 34 maytacatct tcacaacmah wtcaaytca ct	32
40	<210> 35 <211> 24 <212> DNA <213> bacteriano	
45	<400> 35 cmatayyaaa ytacaataaa actc	24
50	<210> 36 <211> 24 <212> DNA <213> bacteriano	
55	<400> 36 caataymaaa ctayaataaa aatt	24
60	<210> 37 <211> 31 <212> DNA <213> bacteriano	
	<400> 37 ggtgaartr tartrkwgt gaagatgdk g	31
60	<210> 38 <211> 31 <212> DNA <213> bacteriano	
	<400> 38 agtgaarttg awdtkgtgt gaagatgtar t	31
	<210> 39 <211> 25	

ES 2 399 054 T3

	<212> DNA <213> bacteriano	
5	<400> 39 gatwggatgg aaagatttr trgag	25
10	<210> 40 <211> 25 <212> DNA <213> bacteriano	
15	<400> 40 kgtwagatgg aaagatttg tgaat	25
20	<210> 41 <211> 23 <212> DNA <213> bacteriano	
25	<400> 41 hymaymmway haaaataata tcc	23
30	<210> 42 <211> 23 <212> DNA <213> bacteriano	
35	<400> 42 tcaammmywm maaaataata ttt	23
40	<210> 43 <211> 26 <212> DNA <213> bacteriano	
45	<400> 43 awccattcta aaaaaacctt taaaca	26
50	<210> 44 <211> 26 <212> DNA <213> bacteriano	
55	<400> 44 aaccawwmyw aamhmacctt cawact	26
60	<210> 45 <211> 28 <212> DNA <213> bacteriano	
65	<400> 45 gttgtaagg tgatatgaat tgtataa	28
70	<210> 46 <211> 26 <212> DNA <213> bacteriano	
75	<400> 46 ttattattaa tgaatttat aggta	26
80	<210> 47 <211> 27	

# ES 2 399 054 T3

	<212> DNA <213> bacteriano	
5	<400> 47 gaggagttta gagtttgaat tagtrtg	27
10	<210> 48 <211> 26 <212> DNA <213> bacteriano	
15	<400> 48 tatatacaaa actatcaccc tatatc	26
20	<210> 49 <211> 22 <212> DNA <213> bacteriano	
25	<400> 49 tcatcaaact cacaacayat ac	22
30	<210> 50 <211> 25 <212> DNA <213> bacteriano	
35	<400> 50 ttgagtaaga tattgatggg ggtaa	25
40	<210> 51 <211> 21 <212> DNA <213> bacteriano	
45	<400> 51 tatggtagg gggttattgt a	21
50	<210> 52 <211> 24 <212> DNA <213> bacteriano	
55	<400> 52 aatctatcat ttaaacctt aacc	24
60	<210> 53 <211> 24 <212> DNA <213> bacteriano	
65	<400> 53 cctaactatc tataccttcc cact	24
70	<210> 54 <211> 25 <212> DNA <213> bacteriano	
75	<400> 54 cactccccta ccataccaat aaacc	25
80	<210> 55 <211> 28	

ES 2 399 054 T3

	<212> DNA <213> bacteriano	
5	<400> 55 gtatgatgag ttagggagtt aagttaa	28
10	<210> 56 <211> 21 <212> DNA <213> bacteriano	
15	<400> 56 ggtgaggta agggatat a	21
20	<210> 57 <211> 30 <212> DNA <213> bacteriano	
25	<400> 57 aaaagagtga agagtggtt ggttagata	30
30	<210> 58 <211> 24 <212> DNA <213> bacteriano	
35	<400> 58 tccaaacctt ttcaacatt aact	24
40	<210> 59 <211> 28 <212> DNA <213> bacteriano	
45	<400> 59 ccctaaaatt attcaaaaa aaacaaaa	28
50	<210> 60 <211> 30 <212> DNA <213> bacteriano	
55	<400> 60 ttagtggggg ttattggtt tattaatgga	30
60	<210> 61 <211> 30 <212> DNA <213> bacteriano	
65	<400> 61 taaggaagtg atgattgaa gatagtgga	30
70	<210> 62 <211> 21 <212> DNA <213> bacteriano	
75	<400> 62 acaccttct tactaaatac t	21
80	<210> 63 <211> 25	

ES 2 399 054 T3

	<212> DNA <213> bacteriano	
5	<400> 63 tataccataa atcttcacta atatc	25
10	<210> 64 <211> 27 <212> DNA <213> bacteriano	
15	<400> 64 ttgtgtagat gatggagtag taggta	27
20	<210> 65 <211> 28 <212> DNA <213> bacteriano	
25	<400> 65 gaatgatgga gtaagttaag tatgtgga	28
30	<210> 66 <211> 27 <212> DNA <213> bacteriano	
35	<400> 66 taaaaattat ttcttaaaaa cctcact	27
40	<210> 67 <211> 26 <212> DNA <213> bacteriano	
45	<400> 67 aaattatctc acacaccta aaatat	26
50	<210> 68 <211> 24 <212> DNA <213> bacteriano	
55	<400> 68 aatgttaaaa ggtaaaggg atat	24
60	<210> 69 <211> 28 <212> DNA <213> bacteriano	
65	<400> 69 tattgaattt aagtttgggt gaatgggt	28
70	<210> 70 <211> 29 <212> DNA <213> bacteriano	
75	<400> 70 ccaatattc aacattaact ccactctc	29
	<210> 71 <211> 29	

ES 2 399 054 T3

	<212> DNA <213> bacteriano	
5	<400> 71 atatccatct tccaaattca taaaataat	29
10	<210> 72 <211> 24 <212> DNA <213> bacteriano	
15	<400> 72 taaacaacaa caattccact ttcc	24
20	<210> 73 <211> 28 <212> DNA <213> bacteriano	
25	<400> 73 ataggaaaag aaawtgaawg wgattttg	28
30	<210> 74 <211> 28 <212> DNA <213> bacteriano	
35	<400> 74 gtgtagtggt gagtgaaggt ggaatagg	28
40	<210> 75 <211> 32 <212> DNA <213> bacteriano	
45	<400> 75 taaacaamtt cmmtcaaaat aacatttlyc aa	32
50	<210> 76 <211> 23 <212> DNA <213> bacteriano	
55	<400> 76 ctaattaata tttaaaactta ccc	23
60	<210> 77 <211> 25 <212> DNA <213> bacteriano	
65	<400> 77 tttgaaatt atatgttat aatgt	25
70	<210> 78 <211> 27 <212> DNA <213> bacteriano	
75	<400> 78 aagtatgagt tggtaggta tgatagt	27
80	<210> 79 <211> 23	

# ES 2 399 054 T3

	<212> DNA <213> bacteriano	
5	<400> 79 cctccamtta wtyataatct yac	23
10	<210> 80 <211> 25 <212> DNA <213> bacteriano	
15	<400> 80 caccwaaaya acaccatcat acatt	25
20	<210> 81 <211> 29 <212> DNA <213> bacteriano	
25	<400> 81 tgtagftaga tagtggggta taagtttta	29
30	<210> 82 <211> 24 <212> DNA <213> bacteriano	
35	<400> 82 aggggaagag ttagattat taaa	24
40	<210> 83 <211> 28 <212> DNA <213> bacteriano	
45	<400> 83 ataactcaw cymwataca acactcat	28
50	<210> 84 <211> 28 <212> DNA <213> bacteriano	
55	<400> 84 atcaatttaa aaaatttca ctycaaaa	28
60	<210> 85 <211> 26 <212> DNA <213> bacteriano	
65	<400> 85 ttttatwat tgatttggg gwtaaa	26
70	<210> 86 <211> 22 <212> DNA <213> bacteriano	
75	<400> 86 tkktwwttag tattgagaat ga	22
80	<210> 87 <211> 24	

ES 2 399 054 T3

	<212> DNA <213> bacteriano	
5	<400> 87 tgtaaattwa tttgtaagt twgt	24
10	<210> 88 <211> 23 <212> DNA <213> bacteriano	
15	<400> 88 gaatgagggg ggattgtta att	23
20	<210> 89 <211> 26 <212> DNA , <213> bacteriano	
25	<400> 89 tctataacca aaacaatcaa aaaata	26
30	<210> 90 <211> 26 <212> DNA <213> bacteriano	
35	<400> 90 cattacacct aacaaatc tcacc	26
40	<210> 91 <211> 26 <212> DNA <213> bacteriano	
45	<400> 91 atwwataggt tgaataggtr agaaat	26
50	<210> 92 <211> 28 <212> DNA <213> bacteriano	
55	<400> 92 atagtgattt ggtggttag tatggaat	28
60	<210> 93 <211> 32 <212> DNA <213> bacteriano	
65	<400> 93 caaacctact tcaactcaaa aataaaataa at	32
70	<210> 94 <211> 29 <212> DNA <213> bacteriano	
75	<400> 94 acaacaattt aaaccaact cacatatct	29
80	<210> 95 <211> 29	

ES 2 399 054 T3

	<212> DNA <213> bacteriano	
5	<400> 95 aaaayaamwv tyttcaatct tcctayaaa	29
10	<210> 96 <211> 27 <212> DNA <213> bacteriano	
15	<400> 96 atwwttgta aggdwrtgar raggaag	27
20	<210> 97 <211> 24 <212> DNA <213> bacteriano	
25	<400> 97 tagragggta aattgargwg tta	24
30	<210> 98 <211> 26 <212> DNA <213> bacteriano	
35	<400> 98 tkattggga arrtwrgta aagaga	26
40	<210> 99 <211> 25 <212> DNA <213> bacteriano	
45	<400> 99 tctctcaac ttaacctcac atcat	25
50	<210> 100 <211> 24 <212> DNA <213> bacteriano	
55	<400> 100 ataattcaa atctacawcm waat	24
60	<210> 101 <211> 25 <212> DNA <213> bacteriano	
65	<400> 101 ratktattgg aggattgaat taggg	25
70	<210> 102 <211> 23 <212> DNA <213> bacteriano	
75	<400> 102 atgttgaaaa gtgtttgat gat	23
80	<210> 103 <211> 31	

	<212> DNA <213> bacteriano	
5	<400> 103 tctaaaatya ataawccaaa ataamcccct.c	31
10	<210> 104 <211> 23 <212> DNA <213> bacteriano	
15	<400> 104 actaccaayh atawhtcatt aac	23
20	<210> 105 <211> 25 <212> DNA <213> bacteriano	
25	<400> 105 aggtgakat tttgtatta gagta	25
30	<210> 106 <211> 29 <212> DNA <213> bacteriano	
35	<400> 106 rwagtgatgg agggatgtag taggtaat	29
40	<210> 107 <211> 25 <212> DNA <213> bacteriano	
45	<400> 107 ctttctyaa caatataaca tcaact	25
50	<210> 108 <211> 24 <212> DNA <213> bacteriano	
55	<400> 108 ctctcamtca cctaaaacta ctca	24
60	<210> 109 <211> 28 <212> DNA <213> bacteriano	
65	<400> 109 agaagttgat gaaggatggt attaatga	28
70	<210> 110 <211> 29 <212> DNA <213> bacteriano	
75	<400> 110 gttatgata tgatgaatwta tagtatrrt	29
80	<210> 111 <211> 24	

ES 2 399 054 T3

	<212> DNA <213> bacteriano	
5	<400> 111 caaaaythtt accttctyta atyc	24
10	<210> 112 <211> 23 <212> DNA <213> bacteriano	
15	<400> 112 caacaaaatt ycacatactc cat	23
20	<210> 113 <211> 23 <212> DNA <213> bacteriano	
25	<400> 113 gatttgatgt aaggtaagt agt	23
30	<210> 114 <211> 31 <212> DNA <213> bacteriano	
35	<400> 114 ttggttaggt tgaagtttag gtaatattga a	31
40	<210> 115 <211> 32 <212> DNA <213> bacteriano	
45	<400> 115 gatttatggt gaaaagtgag tggatgaatt ga	32
50	<210> 116 <211> 29 <212> DNA <213> bacteriano	
55	<400> 116 cctytttcta actcccaaat taaattaat	29
60	<210> 117 <211> 25 <212> DNA <213> bacteriano	
65	<400> 117 gaagtggtgg attgtttttt ggata	25
70	<210> 118 <211> 30 <212> DNA <213> bacteriano	
75	<400> 118 aagggtgtg aagtatgatt gtaaggatat	30
80	<210> 119 <211> 31	

ES 2 399 054 T3

	<212> DNA <213> bacteriano	
5	<400> 119 tacamtccaa ymacacactt cacctatcct a	31
10	<210> 120 <211> 26 <212> DNA <213> bacteriano	
	<400> 120 caacaatata aatcaacaa ctcaaa	26
15	<210> 121 <211> 26 <212> DNA <213> bacteriano	
20	<400> 121 aggagtggft agttttgtg aagtta	26
25	<210> 122 <211> 24 <212> DNA <213> bacteriano	
30	<400> 122 acaaattaa aawccaacac aact	24
35	<210> 123 <211> 26 <212> DNA <213> bacteriano	
	<400> 123 taacactatc tcccaccaya atmaat	26
40	<210> 124 <211> 23 <212> DNA <213> bacteriano	
45	<400> 124 taggtgttg agtttaatt ata	23
50	<210> 125 <211> 24 <212> DNA <213> bacteriano	
	<400> 125 gaagtataaa gtaatggtgg ggtg	24
55	<210> 126 <211> 31 <212> DNA <213> bacteriano	
60	<400> 126 tacaatatca actacaccac ttctaacaaa t	31
	<210> 127 <211> 23	

ES 2 399 054 T3

	<212> DNA <213> bacteriano	
5	<400> 127 taataaaaat aacaattata ttt	23
10	<210> 128 <211> 28 <212> DNA <213> bacteriano	
15	<400> 128 aaggtgtag agtattaagt attttaag	28
20	<210> 129 <211> 24 <212> DNA <213> bacteriano	
25	<400> 129 gttgataatg tattaggggt tgga	24
30	<210> 130 <211> 28 <212> DNA <213> bacteriano	
35	<400> 130 atatggattt gaaagtttag gtaagatg	28
40	<210> 131 <211> 31 <212> DNA <213> bacteriano	
45	<400> 131 tactactaaa tcaacaacaa caatatccac a	31
50	<210> 132 <211> 23 <212> DNA <213> bacteriano	
55	<400> 132 cttaataactt aaaacattaa tct	23
60	<210> 133 <211> 27 <212> DNA <213> bacteriano	
65	<400> 133 gagaataagt aaaaggtggt agttgtg	27
70	<210> 134 <211> 34 <212> DNA <213> bacteriano	
75	<400> 134 gattttatt ggtttattgt tatttgatat tggt	34
	<210> 135 <211> 31	

ES 2 399 054 T3

	<212> DNA <213> bacteriano	
5	<400> 135 caataatca atatcaacac ccaactttt c	31
10	<210> 136 <211> 23 <212> DNA <213> bacteriano	
15	<400> 136 tacacaccac caaacccata tac	23
20	<210> 137 <211> 24 <212> DNA, <213> bacteriano	
25	<400> 137 gaaaataaat agaaagtgt ggtg	24
30	<210> 138 <211> 23 <212> DNA <213> bacteriano	
35	<400> 138 tgttttatt ggatattgtg ttt	23
40	<210> 139 <211> 26 <212> DNA <213> bacteriano	
45	<400> 139 caataacatc tactacacca aaacac	26
50	<210> 140 <211> 25 <212> DNA <213> bacteriano	
55	<400> 140 catattaaac tactcaaat tacco	25
60	<210> 141 <211> 25 <212> DNA <213> bacteriano	
65	<400> 141 tatgtgtttt ggtgaagatt gtta	25
70	<210> 142 <211> 22 <212> DNA <213> bacteriano	
75	<400> 142 tttgatatt gattggggg tg	22
80	<210> 143 <211> 27	

ES 2 399 054 T3

	<212> DNA <213> bacteriano	
5	<400> 143 ggtttgtaa tggggtgat tgttgag	27
10	<210> 144 <211> 20 <212> DNA <213> bacteriano	
15	<400> 144 cataacttac atcaataaaa	20
20	<210> 145 <211> 34 <212> DNA <213> bacteriano	
25	<400> 145 gtaatcaagg tcgtctgaa gactacaaca tggc	34
30	<210> 146 <211> 34 <212> DNA <213> bacteriano	
35	<400> 146 gcaatttagg ccgcctcgaa gattataata tggc	34
40	<210> 147 <211> 34 <212> DNA <213> bacteriano	
45	<400> 147 gtaattaagg ttgtttgaa gattataata tggc	34
50	<210> 148 <211> 34 <212> DNA <213> bacteriano	
55	<400> 148 gtaatttagg ttgtttgaa gattataata tggc	34
60	<210> 149 <211> 16 <212> DNA <213> bacteriano	
65	<400> 149 aggycgycty gaagay	16
70	<210> 150 <211> 16 <212> DNA <213> bacteriano	
75	<400> 150 aggtgtttt gaagat	16
80	<210> 151 <211> 16	

ES 2 399 054 T3

	<212> DNA <213> bacteriano	
5	<400> 151 taactacgga agatca	16
10	<210> 152 <211> 16 <212> DNA <213> bacteriano	
15	<400> 152 taattatgga ágata	16
20	<210> 153 <211> 16 <212> DNA <213> bacteriano	
25	<400> 153 gtaatcaagg tcgtct	16
30	<210> 154 <211> 16 <212> DNA <213> bacteriano	
35	<400> 154 gtaattaagg ttgttt	16
40	<210> 155 <211> 16 <212> DNA <213> bacteriano	
45	<400> 155 gcaatttagg ccgcct	16
50	<210> 156 <211> 16 <212> DNA <213> bacteriano	
55	<400> 156 gtaatttagg ttgttt	16
60	<210> 157 <211> 40 <212> DNA <213> bacteriano	
65	<400> 157 aagctctga aggtgactct aaatacgaag acatcatcat	40
70	<210> 158 <211> 40 <212> DNA <213> bacteriano	
75	<400> 158 aagccctga aggtgacact aaatacgaag acatcgttat	40
80	<210> 159 <211> 40	

ES 2 399 054 T3

<212> DNA  
 <213> bacteriano  
  
 <400> 159  
 5 aagctcttga aggtgactca aaatacgaag atatcatcat 40  
  
 <210> 160  
 <211> 40  
 <212> DNA  
 10 <213> bacteriano  
  
 <400> 160  
 aagctcttga aggtgatact aagtacgaag acatcatcat 40  
  
 <210> 161  
 <211> 40  
 <212> DNA  
 <213> bacteriano  
 15  
  
 <400> 161  
 20 aagctcttga aggtgactct aaatacgaag atatcatcat 40  
  
 <210> 162  
 <211> 40  
 25 <212> DNA  
 <213> bacteriano  
  
 <400> 162  
 30 aagctcttga aggcgataca gcacatgaag atatcatcat 40  
  
 <210> 163  
 <211> 40  
 <212> DNA  
 <213> bacteriano  
 35  
  
 <400> 163  
 aagctcttga aggtgactct aaatacgaag acatcgttat 40  
  
 <210> 164  
 <211> 40  
 <212> DNA  
 <213> bacteriano  
 40  
  
 <400> 164  
 45 aagctcttga aggtgacact cagtacgaag atatcatcat 40  
  
 <210> 165  
 <211> 40  
 <212> DNA  
 50 <213> bacteriano  
  
 <400> 165  
 aagctcttga aggtgattct aaatacgaag acatcatcat 40  
  
 <210> 166  
 <211> 40  
 <212> DNA  
 <213> bacteriano  
 55  
  
 <400> 166  
 60 aagctcttga aggtgattct aaatacgaag acatcatcat 40  
  
 <210> 167  
 <211> 40

<212> DNA  
 <213> bacteriano  
  
 <400> 167  
 5 aagcycctga aggygaywcv vmryaygaag ayatcrtyat 40  
  
 <210> 168  
 <211> 40  
 <212> DNA  
 10 <213> bacteriano  
  
 <400> 168  
 aagttttga aggtgattt aaatatgaag atattattat 40  
  
 <210> 169  
 <211> 40  
 <212> DNA  
 <213> bacteriano  
  
 20 <400> 169  
 aagttttga aggtgatatt aaatatgaag atattgtat 40  
  
 <210> 170  
 <211> 40  
 25 <212> DNA  
 <213> bacteriano  
  
 <400> 170  
 aagttttga aggtgattt aaatatgaag atattattat 40  
  
 <210> 171  
 <211> 40  
 <212> DNA  
 <213> bacteriano  
 35  
 <400> 171  
 aagttttga aggtgatatt aagtatgaag atattattat 40  
  
 <210> 172  
 <211> 40  
 <212> DNA  
 <213> bacteriano  
 40  
 <400> 172  
 aagttttga aggtgattt aaatatgaag atattattat 40  
  
 <210> 173  
 <211> 40  
 <212> DNA  
 50 <213> bacteriano  
  
 <400> 173  
 aagttttga aggtgatata gtatatgaag atattattat 40  
  
 <210> 174  
 <211> 40  
 <212> DNA  
 <213> bacteriano  
  
 60 <400> 174  
 aagttttga aggtgattt aaatatgaag atattgtat 40  
  
 <210> 175  
 <211> 40

# ES 2 399 054 T3

	<212> DNA	
	<213> bacteriano	
	<400> 175	
5	aagttttga aggtgatatt tagtatgaag atattattat	40
	<210> 176	
	<211> 40	
	<212> DNA	
10	<213> bacteriano	
	<400> 176	
	aagttttga aggtgattt aaatatgaag atattattat	40
15	<210> 177	
	<211> 40	
	<212> DNA	
	<213> bacteriano	
20	<400> 177	
	aagttttga aggtgattt aaatatgaag atattattat	40
	<210> 178	
	<211> 40	
25	<212> DNA	
	<213> bacteriano	
	<400> 178	
	aagttttga aggtgatwtw rwrtatgaag atatrrtat	40
30	<210> 179	
	<211> 23	
	<212> DNA	
	<213> bacteriano	
35	<400> 179	
	tacaacgaca ataaaacggt tga	23
	<210> 180	
40	<211> 23	
	<212> DNA	
	<213> bacteriano	
	<400> 180	
45	tataatgata ataaaatggt tga	23
	<210> 181	
	<211> 23	
	<212> DNA	
50	<213> bacteriano	
	<400> 181	
	tacggagata ataaagttgt tga	23
55	<210> 182	
	<211> 23	
	<212> DNA	
	<213> bacteriano	
60	<400> 182	
	tatggagata ataaagttgt tga	23
	<210> 183	
	<211> 23	

ES 2 399 054 T3

	<212> DNA <213> bacteriano	
5	<400> 183 tacaacgaca ataaaacggt tga	23
10	<210> 184 <211> 23 <212> DNA <213> bacteriano	
15	<400> 184 tataatgata ataaaatggt tga	23
20	<210> 185 <211> 23 <212> DNA <213> bacteriano	
25	<400> 185 tacaacgaca ataaaacggt tga	23
30	<210> 186 <211> 23 <212> DNA <213> bacteriano	
35	<400> 186 tataatgata ataaaatggt tga	23
40	<210> 187 <211> 23 <212> DNA <213> bacteriano	
45	<400> 187 tatagagata ataaaacgat taa	23
50	<210> 188 <211> 23 <212> DNA <213> bacteriano	
55	<400> 188 tatagagata ataaaatgat taa	23
60	<210> 189 <211> 23 <212> DNA <213> bacteriano	
65	<400> 189 tacagagata ataaaactat taa	23
70	<210> 190 <211> 23 <212> DNA <213> bacteriano	
75	<400> 190 tatagagata ataaaattat taa	23
80	<210> 191 <211> 23	

ES 2 399 054 T3

	<212> DNA <213> bacteriano	
5	<400> 191 tacaatgaca ataaaatggt tga	23
10	<210> 192 <211> 23 <212> DNA <213> bacteriano	
15	<400> 192 tataatgata ataaaatggt tga	23
20	<210> 193 <211> 23 <212> DNA <213> bacteriano	
25	<400> 193 tayrrhgaya ataaarykrt tra	23
30	<210> 194 <211> 23 <212> DNA <213> bacteriano	
35	<400> 194 tatrrwgata ataaartkrt tra	23
40	<210> 195 <211> 20 <212> DNA <213> viral	
45	<400> 195 tgtgtgtgca gggataattg	20
50	<210> 196 <211> 20 <212> DNA <213> viral	
55	<400> 196 tgtatatgta gggacaattg	20
60	<210> 197 <211> 20 <212> DNA <213> viral	
65	<400> 197 tgtgtttgta gagacaactg	20
70	<210> 198 <211> 20 <212> DNA <213> viral	
75	<400> 198 tgtatatgta gggacaattg	20
80	<210> 199 <211> 20	

ES 2 399 054 T3

	<212> DNA <213> viral	
5	<400> 199 tgtgttgca gagataattg	20
10	<210> 200 <211> 20 <212> DNA <213> viral	
15	<400> 200 tgcattgca gggacaattg	20
20	<210> 201 <211> 20 <212> DNA <213> viral	
25	<400> 201 tgcactgca gggataattg	20
30	<210> 202 <211> 20 <212> DNA <213> viral	
35	<400> 202 tgcgttgcc gagataattg	20
40	<210> 203 <211> 20 <212> DNA <213> viral	
45	<400> 203 tgtgcctgta gagataacag	20
50	<210> 204 <211> 20 <212> DNA <213> viral	
55	<220> <221> característica miscelánea <222> (1)..(10) <223> n	
60	<400> 204 tgyryntgym grgayaaywg	20
65	<210> 205 <211> 20 <212> DNA <213> viral	
70	<400> 205 tgtgtgtgta gggataattg	20
75	<210> 206 <211> 20 <212> DNA <213> viral	

# ES 2 399 054 T3

	<b>&lt;400&gt; 206</b> tgtatatgta gggataattg	20
5	<b>&lt;210&gt; 207</b> <b>&lt;211&gt; 20</b> <b>&lt;212&gt; DNA</b> <b>&lt;213&gt; viral</b>	
10	<b>&lt;400&gt; 207</b> tgtgtttgta gagataattg	20
15	<b>&lt;210&gt; 208</b> <b>&lt;211&gt; 20</b> <b>&lt;212&gt; DNA</b> <b>&lt;213&gt; viral</b>	
20	<b>&lt;400&gt; 208</b> tgtatatgta gggataattg	20
25	<b>&lt;210&gt; 209</b> <b>&lt;211&gt; 20</b> <b>&lt;212&gt; DNA</b> <b>&lt;213&gt; viral</b>	
30	<b>&lt;400&gt; 209</b> tgtgtttgta gagataattg	20
35	<b>&lt;210&gt; 210</b> <b>&lt;211&gt; 20</b> <b>&lt;212&gt; DNA</b> <b>&lt;213&gt; viral</b>	
40	<b>&lt;400&gt; 210</b> tgtatttgta gggataattg	20
45	<b>&lt;210&gt; 211</b> <b>&lt;211&gt; 20</b> <b>&lt;212&gt; DNA</b> <b>&lt;213&gt; viral</b>	
50	<b>&lt;400&gt; 211</b> tgtatttgta gggataattg	20
55	<b>&lt;210&gt; 212</b> <b>&lt;211&gt; 20</b> <b>&lt;212&gt; DNA</b> <b>&lt;213&gt; viral</b>	
60	<b>&lt;400&gt; 212</b> tgtgtttggt gagataattg	20
65	<b>&lt;210&gt; 213</b> <b>&lt;211&gt; 20</b> <b>&lt;212&gt; DNA</b> <b>&lt;213&gt; viral</b>	
70	<b>&lt;400&gt; 213</b> tgtgtttgta gagataatag	20
75	<b>&lt;210&gt; 214</b> <b>&lt;211&gt; 20</b> <b>&lt;212&gt; DNA</b> <b>&lt;213&gt; viral</b>	

ES 2 399 054 T3

	<400> 214 tgtrtdtgtw grgataatwg	20
5	<210> 215 <211> 19 <212> DNA <213> viral	
10	<400> 215 ctaaattcgc tccgattta	19
15	<210> 216 <211> 19 <212> DNA <213> viral	
20	<400> 216 ttaaattgt ttgattta	19
25	<210> 217 <211> 19 <212> DNA <213> viral	
30	<400> 217 caaaattgac ccagactta	19
35	<210> 218 <211> 19 <212> DNA <213> viral	
40	<400> 218 taaaattgat ttagattta	19
45	<210> 219 <211> 19 <212> DNA <213> viral	
50	<400> 219 ttaaattcgt taagattca	19
55	<210> 220 <211> 19 <212> DNA <213> viral	
60	<400> 220 ttaaattgt taagattta	19
	<210> 221 <211> 19 <212> DNA <213> viral	
	<400> 221 ywaaattsry ymmgaytya	19
60	<210> 222 <211> 19 <212> DNA <213> viral	

ES 2 399 054 T3

<400> 222  
twaaatkrk twwgattta

19

5 <210> 223  
<211> 2922  
<212> DNA  
<213> bacteriano

<400> 223  
gattaagtta ttaagggcgc acggtggatg ccttggcact agaagccgat gaaggacgtt 60  
actaacgacg atatgctttg ggtagctgta agtaagcgtt gatccagaga tttccgaatg 120  
ggggaacca gcatgagtta tgtcatgtta tcgatatgtg aatttatagc atgtcagaag 180  
gcagaccggg agaactgaaa catcttagta cccggaggaa gagaaagaaa aatcgattcc 240  
ctgagtagcg gcgagcgaag cgggaagagc ccaaaccaac aagcttgctt gttggggttg 300  
taggacactc tatacggagt tacaaaagaa catgttagac gaatcatctg gaaagatgaa 360  
tcaaagaagg taataatcct gtatcgaag acatattctc tcttgagtgg atcctgagta 420  
cgacggagca cgtgaaatc cgtcggaaac tgggaggacc atctcctaag gctaaatact 480  
ctctagtgac cgatagtgaa ccagtaccgt gagggaaaagg tgaaaagtac cccggaaggg 540  
gagtgaaga gaactgaaa ccgtgtgctt acaagtagtc agagcccgtt aatgggtgat 600  
ggcgtgcctt ttgtagaatg aaccggcgag ttacgatctg atgcaaggtt aagcagcaaa 660  
tgcggagccg cagcgaagc gactctgaat agggcgttga gtatttggc gtagaccga 720  
aaccaggtga tctacccttg gtcaggttga agttcaggta aactgaatg gaggaccgaa 780  
ccgacttacg ttgaaaagtg agcggatgaa ctgagggtag cggagaaat ccaatcgaac 840  
ttggagatag ctggttctct ccgaaatagc tttagggcta gcctcaagtg atgattattg 900  
gaggtagagc actgtttggg cgaggggccc ctctcgggtt accgaattca gacaaactcc 960  
gaatgccaat taatttaact tgggagtcag aacatgggtg ataaggccg tgttcgaaag 1020  
ggaaacagcc cagaccacca gctaaggctc caaaatataat gttaagtgga aaaggatgtg 1080  
gcgttgcca gacaactagg atgttggtt agaagcagcc atcatttaa gagtgcgtaa 1140  
tagctcacta gtcgagtgac actgcgccga aatgtaccg gggctaaaca tattaccgaa 1200  
gctgtggatt gtcctttgga caatggtagg agagcgttct aaggcgtcg aagcatgatc 1260  
gcaaggacat gtggagcgt tagaagtgag aatgccgtg tgagtagcga aagacgggtg 1320  
agaatcccgt ccaccgatg actaaggtt ccagaggaag gctcgtccg tctgggttag 1380  
tcgggtccta agctgaggcc gacaggcgta ggcgatgat aacaggtg taattcctgta 1440  
ccacctagta tcgttttaat cgatgggggg acgcagtagg ataggcgaag cgtgctgttg 1500  
gagtgcacgt ccaagcagta aggctgagt ttaggcaaat ccggcactca taaggctgag 1560  
ctgtgatggg gagaggaaat tgtttcctcg agtcgttgat ttcacactgc cgagaaaagc 1620  
ctctagatag ataacaggtg cccgtaccgc aaaccgacac aggtagtcaa gatgagaatt 1680  
ctaagggtgag cgagcgaact ctctgtaagg aactcggcaa aatgacccc taacttcggg 1740  
agaaggggtg ctctttaggg ttcacgcca gaagagccg agtgaatagg cccaagcgac 1800  
tgtttatcaa aaacacaggt ctctgctaaa ccgtaagggtg atgtatagg gctgacgcct 1860  
gccccgtgct ggaaggttaa gaggagtgg tagcttctgc gaagctacga atcgaagccc 1920  
cagtaaaccg cggccgtaac tataacggtc ctaaggtagc gaaattcctt gtcgggtaa 1980  
10 ttccgaccg cacgaaaggc gtaacgattt gggcactgtc tcaacgagag actcgggtgaa 2040

ES 2 399 054 T3

atcatagtac ctgtgaagat gcaggttacc cgcgacagga cggaaagacc ccgtaggagct 2100  
 ttactgtagc ctgatattga aattcggcac agcttgtaca ggataggtag gagcctttga 2160  
 aacgtgagcg ctagcttacg tggaggcggt ggtgggatac taccctagct gtggtggcctt 2220  
 tctaaccgcc accacttacc gtggtgggag acagtgtcag gcgggcagtt tgactggggc 2280  
 ggtcgcctcc taaaaggtaa cggaggcgct caaaggttcc ctcagaatgg ttggaaatca 2340  
 ttcatagagt gtaaaggcat aaggagcctt gactgcgaga cctacaagtc gagcagggtc 2400  
 gaaagacgga cttagtgatc cggtggttcc gcatggaagg gccatcgctc aacggataaa 2460  
 agctaccccg gggataacag gcttatctcc cccaagagtt cacatcgacg gggaggtttg 2520  
 gcacctcgat gtcggctcat cgcacctcgg ggctgtagtc ggtcccaagg gttgggctgt 2580  
 tcgccatta aagcggtagc cgagctgggt tcagaacgtc gtgagacagt tcggtcccta 2640  
 tccgtcgtgg gcgtaggaaa tttgagagga gctgtcctta gtacgagagg accgggatgg 2700  
 acatacctct ggtgtaccag ttgtcgtgcc aacggcatag ctgggtagct atgtatggac 2760  
 gggataagtg ctgaaagcat ctaagcatga agccccctc aagatgagat ttcccaactt 2820  
 cggttataag atccctcgaa gatgacgagg ttaataggtt cgagggtgaa gcgtggtgac 2880  
 acgtggagct gacgaatact aatcgatcga agacttaatc aa 2922

<210> 224  
 <211> 2922  
 <212> DNA  
 <213> bacteriano

5

<400> 224  
 gattaagtta ttaaggggtg atgggtgatg ttttggtatt agaagttgat gaaggatggt 60  
 attaatgatg atatgttttg ggtagttgta agtaagtgtt galtttagaga tttttgaatg 120  
 ggggaattta gtatgagtta tgttatgtta ttgatatgtg aatttatagt atgttagaag 180  
 gtagatttgg agaattgaaa tatttttagta tttggaggaa gagaaagaaa aattgatttt 240  
 ttgagtagtg gtgagtgaaa tgggaagagt ttaaattaat aagtttgttt gttgggggtg 300  
 taggatattt tatatggagt tataaaagaa tatgttagat gaattatttg gaaagatgaa 360  
 ttaaagaagg taataatttt gtagttgaaa atatatTTTT ttttgagtgg attttgagta 420  
 tgatggagta tgtgaaattt tgttggaaatt tgggaggatt attttttaag gttaaattt 480  
 ttttagtgat tgatagttaa ttagtattgt gagggaaagg tgaaaagtat tttggaaggg 540  
 gagtgaaaga gaatttgaag ttgtgtgttt ataagtagtt agagtttgtt aatgggtgat 600  
 ggtgtgtttt ttgtagaatg aattgggtgag ttatgatttg atgtaagggt aagtagtaaa 660  
 tgtggagtgg tagtgaaagt gagtttgaat aggggtgtga gtatttgggt gtagatttga 720  
 aattaggtga tttatttttg gttaggttga agtttaggta atattgaatg gaggattgaa 780  
 ttgatttatg ttgaaaagtg agtggatgaa ttgagggtag tggagaaatt ttaattgaat 840  
 ttggagatag ttggtttttt ttgaaatagt tttagggtta gttttaagtg atgattattg 900  
 gaggtagagt attgtttgga tgaggggttt tttttgggtt attgaattta gataaatttt 960  
 gaatgttaat taatttaatt tgggagttag aatatgggtg ataaggtttg tgtttgaaag 1020  
 ggaaatagtt tagattatta gttaaggttt taaaatatat gttaagtgga aaaggatgtg 1080  
 gtgttgttta gataattagg atgttggttt agaagtagtt attatttaa gagtgtgtaa 1140  
 tagtttatta gttgagtgat attgtgttga aaatgtattg gggtaaata tattattgaa 1200  
 gttgtggatt gttttttgga taatggtagg agagtgtttt aagggtgttg aagtatgatt 1260  
 gtaaggatat gtggagtgtt tagaagtgag aatgttgggt tgagtagtga aagatgggtg 1320

ES 2 399 054 T3

```

agaat t t t t t t g t   t t a t t g a t t g   a t t a a g g t t t   t t a g a g g a a g   g t t t g t t t g t   t t t g g g t t a g   1380
t t g g g t t t t a   a g t t g a g g t t   g a t a g g t g t a   g g t g a t g g a t   a a t a g g t t g a   t a t t t t t g t a   1440
t t a t t t a g t a   t t g t t t t a a t   t g a t g g g g g g   a t g t a g t a g g   a t a g g t g a a g   t g t g t t g t t g   1500
g a g t g t a t g t   t t a a g t a g t a   a g g t t g a g t g   t t a g g t a a a t   t t g g t a t t t a   t a a g g t t g a g   1560
t t g t g a t g g g   g a g a g g a a a t   t g t t t t t t t g   a g t t g t t g a t   t t t a t a t t g t   t g a g a a a a g t   1620
t t t t a g a t a g   a t a a t a g g t g   t t t g t a t t g t   a a a t t g a t a t   a g g t a g t t a a   g a t g a g a a t t   1680
t t a a g g t g a g   t g a g t g a a t t   t t t g t t a a g g   a a t t t g g t a a   a a t g a t t t t g   t a a t t t t g g g   1740
a g a a g g g g t g   t t t t t t a g g g   t t t a t g t t t a   g a a g a g t t g t   a g t g a a t a g g   t t t a a g t g a t   1800
t g t t t a t t a a   a a a t a t a g g t   t t t t g t t a a a   t t g t a a g g t g   a t g t a t a g g g   g t t g a t g t t t   1860
g t t t g g t g t t   g g a a g g t t a a   g a g g a g t g g t   t a g t t t t t g t   g a a g t t a t g a   a t t g a a g t t t   1920
t a g t a a a t g g   t g g t t g t a a t   t a t a a t g g t t   t t a a g g t a g t   g a a a t t t t t t   g t t g g g t a a g   1980
t t t t g a t t t g   t a t g a a a g g t   g t a a t g a t t t   g g g t a t t g t t   t t a a t g a g a g   a t t t g g t g a a   2040
a t t a t a g t a t   t t g t g a a g a t   g t a g g t t a t t   t g t g a t a g g a   t g g a a a g a t t   t t g t g g a g t t   2100
t t a t t g t a g t   t t g a t a t t g a   a a t t t g g t a t   a g t t t g t a t a   g g a t a g g t a g   g a g t t t t t g a   2160
a a t g t g a g t g   t t a g t t t a t g   t g g a g g t g t t   g g t g g g a t a t   t a t t t t a g t t   g t g t t g g t t t   2220
t t t a a t t t g t   a t t a t t t a t t   g t g g t g g g a g   a t a g t g t t a g   g t g g g t a g t t   t g a t t g g g g t   2280
g g t t g t t t t t   t a a a a g g t a a   t g g a g g t g t t   t a a a g g t t t t   t t t a g a a t g g   t t g g a a a t t a   2340
t t t a t a g a g t   g t a a a g g t a t   a a g g g a g t t t   g a t t g t g a g a   t t t a t a a g t t   g a g t a g g g t t   2400
g a a a g a t g g a   t t t a g t g a t t   t g g t g g t t t t   g t a t g g a a g g   g t t a t t g t t t   a a t g g a t a a a   2460
a g t t a t t t t g   g g g a t a a t a g   g t t t a t t t t t   t t t a a g a g t t   t a t a t t g a t g   g g g a g g t t t g   2520
g t a t t t t g a t   g t t g g t t t a t   t g t a t t t t g g   g g t t g t a g t t   g g t t t t a a g g   g t t g g g t t g t   2580
t t g t t t a t t a   a a g t g g t a t g   t g a g t t g g g t   t t a g a a t g t t   g t g a g a t a g t   t t g g t t t t t a   2640
t t t g t t g t g g   g t g t a g g a a a   t t t g a g a g g a   g t t g t t t t t a   g t a t g a g a g g   a t t g g g a t g g   2700
a t a t a t t t t t   g g t g t a t t a g   t t g t t g t g t t   a a t g g t a t a g   t t g g g t a g t t   a t g t a t g g a t   2760
g g g a t a a g t g   t t g a a a g t a t   t t a a g t a t g a   a g t t t t t t t t   a a g a t g a g a t   t t t t t a a t t t   2820
t g g t t a t a a g   a t t t t t t g a a   g a t g a t g a g g   t t a a t a g g t t   t g a g g t g g a a   g t g t g g t g a t   2880
a t g t g g a g t t   g a t g a a t a t t   a a t t g a t t g a   a g a t t t a a t t   a a

```

- <210> 225
- 5 <211> 1062
- <212> DNA
- <213> bacteriano
  
- <400> 225

ES 2 399 054 T3

atggctatcg acgaaaacaa acagaaagcg ttggcggcag cactgggcca gattgagaaa 60  
caat ttggta aaggctccat catgcgctg ggtgaagacc gttccatgga tgtggaaacc 120  
atctctaccg gtctgctttc actggatabc gcgcttgggg caggtggctc gccgatgggc 180  
cgtatcgtcg aaatctacgg accggaatct tccggtaaaa ccacgctgac gctgcaggtg 240  
atcgccgcag cgcagcgtga aggtaaaacc tgtgctgta tcgatgctga acacgcgctg 300  
gaccaaatct acgcacgtaa actgggctc gatatcgata acctgctgtg ctcccagccg 360  
gacaccggcg agcaggcact ggaaatctgt gacgccctgg cgcgttctgg cgcagtagac 420  
gttatcgtcg ttgactccgt ggcggcactg acgccgaaag cggaaatcga aggcgaaatc 480  
ggcgactctc acatgggctt tgcggcacgt atgatgagcc aggcgatgag taagctggcg 540  
ggtaacctga agcagtccaa cacgctgctg atcttcatca accagatccg tatgaaaatt 600  
ggtgtgatgt tcggtaacct ggaaaccact accggtggta acgcgctgaa attctacgcc 660  
tctgttcgtc tcgacatccg tcgtatcggc gcggtgaaag agggcgaaaa cgtgggtgggt 720  
agcgaaaccg cgtgaaagt ggtgaagaac aaaatcgtc gcgccgttaa acaggctgaa 780  
ttccagatcc tctacggcga aggtatcaac ttctacggcg aactggttga cctgggcgta 840  
aaagagaagc tgatcgagaa agcaggcgcg tggtagcgt acaaaggtag gaagatcggc 900  
cagggtaaag cgaatgcgac tgcctggctg aaagataacc cggaaaccgc gaaagagatc 960  
gagaagaaag tacgtgagtt gctgctgagc aaccggaact caacgccgga tttctctgta 1020  
gatgatagcg aaggcgtagc agaaactaac gaagat tttt aa 1062

<210> 226  
<211> 1062  
<212> DNA  
<213> bacteriano

5

<400> 226  
atggttattg atgaaaataa atagaaagtg ttgggtggtag tattgggtta gattgagaaa 60  
taat ttggta aagg ttttat tatgtg tttg ggtgaagatt gttttatgga tgtggaaatt 120  
at ttttattg gtttg ttttt attggatatt gtgtttgggg taggtggttt gttgatgggt 180  
tgtattgttg aaat tttatg attggaattt tttggtaaaa ttatgttgat gttgtaggtg 240  
attgtttag tagtagtgga aggtaaaatt tgtgtgttta ttgatgttga atatgtgttg 300  
gatttaattt atgtatgtaa attgggtgtt gatattgata atttgttctg ttttagttg 360  
gatattgggtg agtaggtatt ggaaatttgt gatgttttgg tgtgttttgg ttagtagat 420  
gttattgttg ttgattttgt ggtggattg atgttgaaag tggaaattga aggtgaaatt 480  
gggtgattttt atatgggttt tgtggatgt atgatgagtt aggtgatgtg taagttgggtg 540  
ggtaatttga agtagtttaa tatgttgttg atttttatta attagatttg tatgaaaatt 600  
ggtgtgatgt ttggtaattt ggaaattatt attgggtgga atgtgttgaa attttatgtt 660  
tttgtttgtt ttgatatttg ttgtattgg gtgggtgaaag aggggtgaaaa tgtgggtgggt 720  
agtgaaattt gtgtgaaagt ggtgaagaat aaaattgttg tgttgtttaa ataggttgaa 780  
ttttagattt tttatggtag aggtat tttatgggtg aattggttga tttgggtgta 840  
aaagagaagt tgattgagaa agtaggtgtg tggtagatg ataaaggtag gaagattggc 900  
tagggtaaaag tgaatgtgat tgtttgggtg aaagataatt tggaaattgt gaaagagatt 960  
gagaagaaag tatgtgagtt gttgttgagt aatttgaatt taatgttggga tttttttgta 1020  
gatgatagtg aagggttagt agaaattaat gaagat tttt aa 1062

10

## REIVINDICACIONES

1. Un método para detectar un microorganismo que comprende:
- 5 reducir la complejidad de un genoma microbiano o de un ácido nucleico microbiano mediante la generación de una forma derivada del genoma microbiano o de una forma derivada del ácido nucleico microbiano en el que las posiciones ocupadas naturalmente por citosinas están ocupadas por uracilos en la forma derivada del genoma microbiano o forma derivada del ácido nucleico microbiano mediante el tratamiento del genoma microbiano o un ácido nucleico microbiano con un agente seleccionado entre bisulfito, acetato o citrato que modifica la citosina a uracilo;
- 10 donde la forma derivada del genoma microbiano o la forma derivada del ácido nucleico microbiano contienen un ácido nucleico que es específico para un microorganismo que tiene el genoma microbiano o el ácido nucleico microbiano, y
- detectar el ácido nucleico específico microbiano, en donde la detección del ácido nucleico específico microbiano es indicativa de la presencia de un microorganismo que tiene el genoma microbiano o el ácido nucleico microbiano.
- 15 2. El método de acuerdo con la reivindicación 1, en donde los uracilos en la forma derivada del genoma microbiano o la forma derivada del ácido nucleico microbiano son remplazados por timinas mediante la amplificación del genoma microbiano derivado o del ácido nucleico microbiano derivado para formar un genoma microbiano simplificado o un ácido nucleico microbiano simplificado que contiene un ácido nucleico que es específico para un microorganismo que tiene el genoma microbiano o el ácido nucleico microbiano.
- 20 3. El método de acuerdo con la reivindicación 1 o 2, que comprende convertir el ARN microbiano en ADN antes de llevar a cabo el método.
4. El método de acuerdo con la reivindicación 1 o 2, que comprende llevar a cabo el método sobre ARN microbiano para producir una molécula de ARN derivada o simplificada convirtiendo a continuación el ARN derivado o simplificado para formar una molécula de ADN derivada o simplificada.
- 25 5. El método de acuerdo con cualquiera de las reivindicaciones 1 a 4, en donde el agente es bisulfito de sodio.
6. El método de acuerdo con una cualquiera de las reivindicaciones 1 a 5, en donde el ácido nucleico específico microbiano comprende una o más secuencias de nucleótidos únicas para un microorganismo.
7. El método de acuerdo con una cualquiera de las reivindicaciones 2 a 6, en donde el ácido nucleico específico microbiano contiene sustancialmente solo bases seleccionadas entre adenina (A), guanina (G) y timina (T).
- 30 8. El método de acuerdo con una cualquiera de las reivindicaciones 1 a 7, en donde el genoma microbiano o el ácido nucleico es de fago, virus, viroide, bacteria, hongo, alga, protozoo, espiroqueta, u organismo unicelular.
9. El método de acuerdo con una cualquiera de las reivindicaciones 1 a 7, en donde el ácido nucleico específico microbiano es de una región de un gen ribosomal de un procariota o de un microorganismo eucariótico unicelular.
- 35 10. El método de acuerdo con la reivindicación 9, en donde la región del gen ribosomal es 16S o 23S en un procariota o 18S o 28S en un microorganismo eucariótico unicelular.
11. El método de acuerdo con cualquiera de las reivindicaciones 2 a 10, en donde la amplificación se lleva a cabo mediante reacción en cadena de la polimerasa (PCR), amplificación isotérmica, o amplificación de la señal.
12. El método de acuerdo con una cualquiera de las reivindicaciones 1 a 10, en donde el ácido nucleico específico microbiano se detecta mediante PCR en tiempo real.
- 40 13. El método de acuerdo con cualquiera de las reivindicaciones 1 a 10, en donde el ácido nucleico específico microbiano es detectado por un sistema de detección de micromatrices.
14. El método de acuerdo con una cualquiera de las reivindicaciones 1 a 10, en donde el ácido nucleico específico microbiano es detectado:
- 45 proporcionando un ligando detector capaz de unirse a una región diana de la molécula de ácido nucleico específica microbiana y dejando tiempo suficiente para que el ligando detector se una a la región diana; y
- midiendo la unión del ligando detector a la región diana para detectar la presencia del ácido nucleico específico microbiano.

15. El método de acuerdo con la reivindicación 11, en donde el ácido nucleico específico microbiano es detectado mediante la separación de un producto de amplificación y la visualización del producto separado.
16. El método de acuerdo con la reivindicación 14, en donde el producto de amplificación se separa mediante electroforesis y se detecta mediante la visualización de una o más bandas en un gel.
- 5 17. El método de acuerdo con una cualquiera de las reivindicaciones 1 a 13, en donde el ácido nucleico específico microbiano no está presente naturalmente en el microorganismo.
18. El método de acuerdo con una cualquiera de las reivindicaciones 1 a 14, en donde la molécula de ácido nucleico específico microbiano tiene una secuencia de ácido nucleico indicativa de un nivel taxonómico del microorganismo.
- 10 19. El método de acuerdo con la reivindicación 18, en donde el nivel taxonómico del microorganismo incluye familia, género, especie, cepa, tipo, o diferentes poblaciones de las mismas o diferentes poblaciones geográficas o bentónicas.
20. Un método para detectar la presencia de un microorganismo, comprendiendo el método:
- 15 tratar ácido nucleico microbiano con un agente seleccionado entre bisulfito, acetato o citrato que modifica la citosina a uracilo para formar un ácido nucleico microbiano derivado en el que las posiciones ocupadas naturalmente por citosinas en el ácido nucleico microbiano están ocupadas por uracilos en el ácido nucleico microbiano derivado;
- proporcionar cebadores capaces de permitir la amplificación de una molécula de ácido nucleico específica microbiana deseada específica de un microorganismo para el ácido nucleico microbiano derivado;
- 20 llevar a cabo una reacción de amplificación sobre el ácido nucleico microbiano derivado para formar un ácido nucleico simplificado que tenga timinas en lugar de uracilos en el ácido nucleico microbiano derivado;
- y
- analizar la presencia de un producto de ácido nucleico amplificado que contiene la molécula de ácido nucleico específica microbiana deseada, en donde la detección del ácido nucleico específico microbiano deseado es indicativa de la presencia del microorganismo.
- 25 21. El método de acuerdo con la reivindicación 20, en donde el microorganismo se selecciona entre fago, virus, viroide, bacteria, hongo, alga, protozoo, espiroqueta, u organismo unicelular.
22. El método de acuerdo con la reivindicación 20 o 21, en donde el agente es bisulfito de sodio.
23. El método de acuerdo con una cualquiera de las reivindicaciones 20 a 22, en donde la amplificación se lleva a cabo mediante reacción en cadena de la polimerasa (PCR), amplificación isotérmica, o amplificación de la señal.
- 30 24. El método de acuerdo con una cualquiera de las reivindicaciones 20 a 23, en donde la molécula de ácido nucleico específica microbiana deseada se detecta mediante PCR en Tiempo Real.
25. El método de acuerdo con una cualquiera de las reivindicaciones 20 a 23, en donde la molécula de ácido nucleico específica microbiana deseada se detecta mediante un sistema de detección de micromatrices.
- 35 26. El método de acuerdo con una cualquiera de las reivindicaciones 20 a 23, en donde la molécula de ácido nucleico específica microbiana deseada es detectada:
- proporcionando un ligando detector capaz de unirse a una región de la molécula de ácido nucleico específica microbiana deseada y dejando tiempo suficiente para que el ligando detector se una a la región; y
- midiendo la unión del ligando detector a la molécula de ácido nucleico específica microbiana deseada para detectar la presencia de la molécula de ácido nucleico.
- 40 27. El método de acuerdo con una cualquiera de las reivindicaciones 20 a 23, en donde la molécula de ácido nucleico específica microbiana deseada se detecta separando un producto de amplificación y visualizando el producto separado.
28. El método de acuerdo con la reivindicación 27, en donde el producto de amplificación se separa mediante electroforesis y se detecta visualizando una o más bandas sobre un gel.

**Figura 1.**

<b>Secuencias de iga de <i>Neisseria</i></b>	
<b>Secuencia No Convertida</b>	
<i>Neisseria meningitidis</i>	GTAATCA AGGTCGTCTT GAAGACTACA ACATGGC (SEQ ID No 145)
<i>Neisseria gonorrhoeae</i>	GCAATTT AGGCCGCCTC GAAGATTATA ATATGGC (SEQ ID No 146)
<b>Secuencia consenso</b>	GYAATYW AGGYCGYCTY GAAGAYTAYA AYATGGC (SEQ ID No 3) 512 Posibles combinaciones cebadores 74% Similitud de secuencia
<b>Secuencia simplificada</b>	
<i>Neisseria meningitidis</i>	GTAATTA AGGTTGTTTT GAAGATTATA ATATGGT (SEQ ID No 147)
<i>Neisseria gonorrhoeae</i>	GTAATTT AGGTTGTTTT GAAGATTATA ATATGGT (SEQ ID No 148)
<b>Secuencia consenso</b>	GTAATTW AGGTTGTTTT GAAGATTATA ATATGGT (SEQ ID No 4) 2 Combinaciones cebadores 97% Similitud de secuencia

**Figura 2.**

<b>Secuencias de iga de <i>Neisseria</i></b>	
<b>Secuencia No Convertida</b>	
<i>Neisseria meningitidis</i>	GTAATCA AGGTCGTCTT GAAGACTACA ACATGGC (SEQ ID No 145)
<i>Neisseria gonorrhoeae</i>	GCAATTT AGGCCGCCTC GAAGATTATA ATATGGC (SEQ ID No 146)
<b>Secuencia INA consenso</b>	AGGYCGYCTY GAAGAY (SEQ ID No 149) 16 Posibles combinaciones cebadores 75% Similitud de secuencia
<b>Secuencia simplificada</b>	
<i>Neisseria meningitidis</i>	GTAATTA AGGTTGTTTT GAAGATTATA ATATGGT (SEQ ID No 147)
<i>Neisseria gonorrhoeae</i>	GTAATTT AGGTTGTTTT GAAGATTATA ATATGGT (SEQ ID No 148)
<b>Secuencia INA consenso</b>	AGGTTGTTTT GAAGAT (SEQ ID No 150) 100% Similitud de secuencia

**Figura 3.**

**Secuencias del gen *iga***

	No convertida	Simplificada
<i>Haemophilus influenza</i>	TAACTACGG AAGATCA(151)	TAATTATGG AAGATTA(152)
<i>Neisseria meningitidis</i>	GTAATCAAG GTCGTCT(153)	GTAATTAAG GTTGTTT(154)
<i>Neisseria gonorrhoeae</i>	GCAATTTAG GCCGCCT(155)	GTAATTTAG GTTGTTT(156)

**Figura 4.**

**Gen *tuf* de *Streptococcus***

**Secuencia No Convertida**

<i>S.oralis</i>	AAGCTCTTGA AGGTGACTCT AAATACGAAG ACATCATCAT (SEQ ID No 157)
<i>S.mitis</i>	AAGCCCTTGA AGGTGACACT AAATACGAAG ACATCGTTAT (SEQ ID No 158)
<i>S.dysgalactiae</i>	AAGCTCTTGA AGGTGACTCA AAATACGAAG ATATCATCAT (SEQ ID No 159)
<i>S.cristatus</i>	AAGCTCTTGA AGGTGATACT AAGTACGAAG ACATCATCAT (SEQ ID No 160)
<i>S.gordonii</i>	AAGCTCTTGA AGGTGACTCT AAATACGAAG ATATCATCAT (SEQ ID No 161)
<i>S.parauberis</i>	AAGCTCTTGA AGGCGATACA GCACATGAAG ATATCATCAT (SEQ ID No 162)
<i>S.pneumoniae</i>	AAGCTCTTGA AGGTGACTCT AAATACGAAG ACATCGTTAT (SEQ ID No 163)
<i>S.bovis</i>	AAGCTCTTGA AGGTGACACT CAGTACGAAG ATATCATCAT (SEQ ID No 164)
<i>S.vestibularis</i>	AAGCTCTTGA AGGTGATTCT AAATACGAAG ACATCATCAT (SEQ ID No 165)
<i>S.uberis</i>	AAGCTCTTGA AGGTGATTCT AAATACGAAG ACATCATCAT (SEQ ID No 166)

Consenso AAGCYCTTGA AGGYGAYWCW VMRYAYGAAG AYATCRTYAT (SEQ ID No 167)  
 67,5% Homología  
 12.288 Posibles combinaciones cebadores

**Secuencia simplificada**

<i>S.oralis</i>	AAGTTTTTGA AGGTGATTTT AAATATGAAG ATATTATTAT (SEQ ID No 168)
<i>S.mitis</i>	AAGTTTTTGA AGGTGATATT AAATATGAAG ATATTGTTAT (SEQ ID No 169)
<i>S.dysgalactiae</i>	AAGTTTTTGA AGGTGATTTA AAATATGAAG ATATTATTAT (SEQ ID No 170)
<i>S.cristatus</i>	AAGTTTTTGA AGGTGATATT AAGTATGAAG ATATTATTAT (SEQ ID No 171)
<i>S.gordonii</i>	AAGTTTTTGA AGGTGATTTT AAATATGAAG ATATTATTAT (SEQ ID No 172)
<i>S.parauberis</i>	AAGTTTTTGA AGGTGATATA GTATATGAAG ATATTATTAT (SEQ ID No 173)
<i>S.pneumoniae</i>	AAGTTTTTGA AGGTGATTTT AAATATGAAG ATATTGTTAT (SEQ ID No 174)
<i>S.bovis</i>	AAGTTTTTGA AGGTGATATT TAGTATGAAG ATATTATTAT (SEQ ID No 175)
<i>S.vestibularis</i>	AAGTTTTTGA AGGTGATTTT AAATATGAAG ATATTATTAT (SEQ ID No 176)
<i>S.uberis</i>	AAGTTTTTGA AGGTGATTTT AAATATGAAG ATATTATTAT (SEQ ID No 177)

Consenso AAGTTTTTGA AGGTGATWTW RWRTATGAAG ATATRTTAT (SEQ ID No 178)  
 85% Homología  
 64 Posibles combinaciones cebadores

**Figura 5.**

**Genes de enterotoxina Estafilocócica (SE)**

	Secuencia No Convertida	Secuencia simplificada
SEC	TAC AACGACAATA AACGGTTGA (179)	TAT AATGATAATA AAATGGTTGA (180)
SEI	TAC GGAGATAATA AAGTTGTTGA (181)	TAT GGAGATAATA AAGTTGTTGA (182)
SEC3	TAC AACGACAATA AACGGTTGA (183)	TAT AATGATAATA AAATGGTTGA (184)
SEC1	TAC AACGACAATA AACGGTTGA (185)	TAT AATGATAATA AAATGGTTGA (186)
SEA	TAT AGAGATAATA AAACGATTAA (187)	TAT AGAGATAATA AAATGATTAA (188)
SEE	TAC AGAGATAATA AAATATTAA (189)	TAT AGAGATAATA AAATATTAA (190)
SEB	TAC AATGACAATA AAATGGTTGA (191)	TAT AATGATAATA AAATGGTTGA (192)
<b>CONSENSO</b>	TAY RRHGAYAATA AARYKRTTRA (193)	TAT RRWGATAATA AARTKRTTRA (194)
	56% Homología	74% Homología
	1536 Combinaciones cebadores	64 Combinaciones cebadores

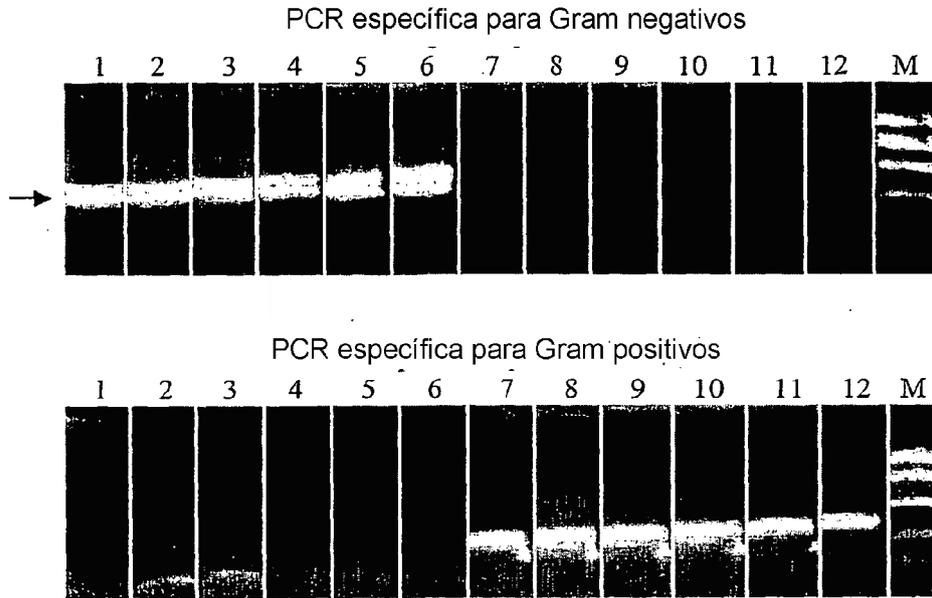
**Figura 6.**

		<b>Neuraminidasa de virus Influenza</b>	
		Secuencia No Convertida	
Virus Influenza A	H5N1	TGTGTGTGCA	GGGATAATTG (SEQ ID No 195)
Virus Influenza A	H7N3	TGTATATGTA	GGGACAATTG (SEQ ID No 196)
Virus Influenza A	H5N8	TGTGTTTGTA	GAGACAAC TG (SEQ ID No 197)
Virus Influenza A	H5N3	TGTATATGTA	GGGACAATTG (SEQ ID No 198)
Virus Influenza A	H5N2	TGTGTTTGCA	GAGATAATTG (SEQ ID No 199)
Virus Influenza A	H6N6	TGCATTTGCA	GGGACAATTG (SEQ ID No 200)
Virus Influenza A	H2N9	TGCACTTGCA	GGGATAATTG (SEQ ID No 201)
Virus Influenza A	H6N5	TGCGTTTGCC	GAGATAATTG (SEQ ID No 202)
Virus Influenza B	NA	TGTGCCTGTA	GAGATAACAG (SEQ ID No 203)
<b>Consenso</b>		TGYRYNTGYM	GRGAYAAYWG (SEQ ID No 204)
		<b>2048 Posibles combinaciones cebadores</b>	
		<b>50% Homología</b>	
		Secuencia simplificada	
Virus Influenza A	H5N1	TGTGTGTGTA	GGGATAATTG (SEQ ID No 205)
Virus Influenza A	H7N3	TGTATATGTA	GGGATAATTG (SEQ ID No 206)
Virus Influenza A	H5N8	TGTGTTTGTA	GAGATAATTG (SEQ ID No 207)
Virus Influenza A	H5N3	TGTATATGTA	GGGATAATTG (SEQ ID No 208)
Virus Influenza A	H5N2	TGTGTTTGTA	GAGATAATTG (SEQ ID No 209)
Virus Influenza A	H6N6	TGTATTTGTA	GGGATAATTG (SEQ ID No 210)
Virus Influenza A	H2N9	TGTATTTGTA	GGGATAATTG (SEQ ID No 211)
Virus Influenza A	H6N5	TGTGTTTGTT	GAGATAATTG (SEQ ID No 212)
Virus Influenza B	NA	TGTGTTTGTA	GAGATAATAG (SEQ ID No 213)
<b>Consenso</b>		TGTRTDGTW	GRGATAATWG (SEQ ID No 214)
		<b>48 Posibles combinaciones cebadores</b>	
		<b>75% Homología</b>	

**Figura 7.**

		<b>Genes VP4 de Rotavirus</b>			
		No convertida	Simplificada		
Rotavirus Cepa	A VP4	CTAAATTCGC	TCCGATTTA (215)	TTAAATTTGT	TTTGATTTA (216)
Rotavirus Cepa	B VP4	CAAATTCGAC	CCAGACTTA (217)	TAAAATTTGAT	TTAGATTTA (218)
Rotavirus Cepa	C VP4	TTAAATTCGT	TAAGATTCA (219)	TTAAATTTGT	TAAGATTTA (220)
<b>Secuencia Consenso</b>		YWAAATTSRY	YMMGAYTYA (221)	TWAAATTKRT	TWWGATTTA (222)
		<b>52% Homología</b>		<b>74% Homología</b>	
		<b>512 Combinaciones cebadores</b>		<b>32 Combinaciones cebadores</b>	

Figura 8.

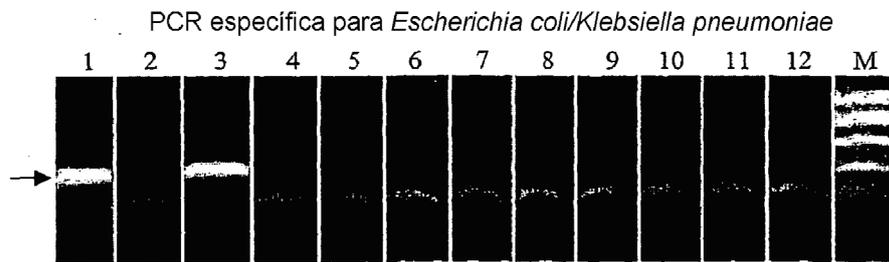


1. *Escherichia coli*
2. *Neisseria gonorrhoeae*
3. *Klebsiella pneumoniae*
4. *Moraxella catarrhalis*
5. *Pseudomonas aeruginosa*
6. *Proteus vulgaris*
7. *Enterococcus faecalis*
8. *Staphylococcus epidermidis*
9. *Staphylococcus aureus*
10. *Staphylococcus xylois*
11. *Streptococcus pneumoniae*
12. *Streptococcus haemolyticus*

**Tinción Gram**

- Negativa
- Negativa
- Negativa
- Negativa
- Negativa
- Negativa
- Positiva
- Positiva
- Positiva
- Positiva
- Positiva
- Positiva

Figura 9.



1. *Escherichia coli*
2. *Neisseria gonorrhoeae*
3. *Klebsiella pneumoniae*
4. *Moraxella catarrhalis*
5. *Pseudomonas aeruginosa*
6. *Proteus vulgaris*
7. *Enterococcus faecalis*
8. *Staphylococcus epidermidis*
9. *Staphylococcus aureus*
10. *Staphylococcus xylois*
11. *Streptococcus pneumoniae*
12. *Streptococcus haemolyticus*

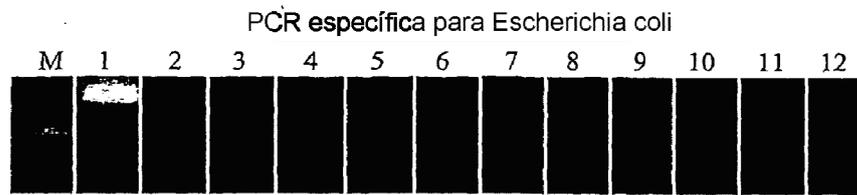
Figura 10.

PCR específica para Neisseria



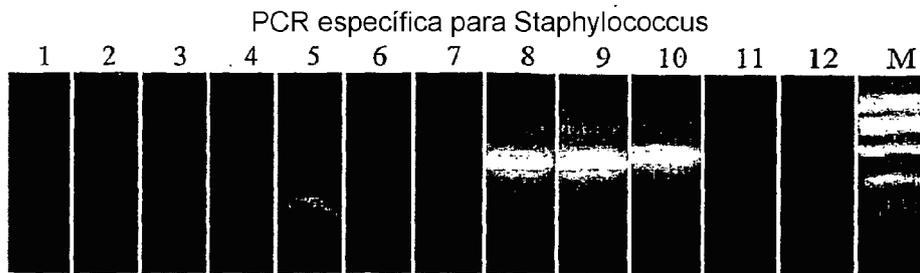
1. *Escherichia coli*
2. *Neisseria gonorrhoeae*
3. *Klebsiella pneumoniae*
4. *Moraxella catarrhalis*
5. *Pseudomonas aeruginosa*
6. *Proteus vulgaris*
7. *Enterococcus faecalis*
8. *Staphylococcus epidermidis*
9. *Staphylococcus aureus*
10. *Staphylococcus xylois*
11. *Streptococcus pneumoniae*
12. *Streptococcus haemolyticus*

Figura 11.



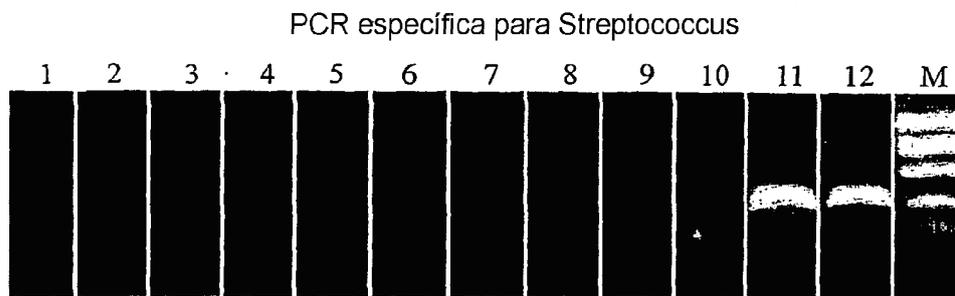
1. *Escherichia coli*
2. *Neisseria gonorrhoeae*
3. *Klebsiella pneumoniae*
4. *Moraxella catarrhalis*
5. *Pseudomonas aeruginosa*
6. *Proteus vulgaris*
7. *Enterococcus faecalis*
8. *Staphylococcus epidermidis*
9. *Staphylococcus aureus*
10. *Staphylococcus xylois*
11. *Streptococcus pneumoniae*
12. *Streptococcus haemolyticus*

Figura 12.



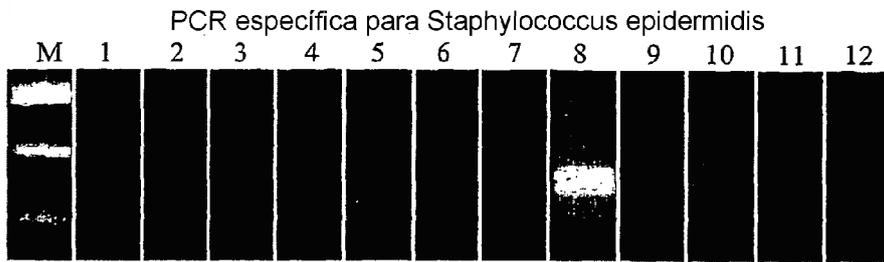
1. *Escherichia coli*
2. *Neisseria gonorrhoeae*
3. *Klebsiella pneumoniae*
4. *Moraxella catarrhalis*
5. *Pseudomonas aeruginosa*
6. *Proteus vulgaris*
7. *Enterococcus faecalis*
8. *Staphylococcus epidermidis*
9. *Staphylococcus aureus*
10. *Staphylococcus xylois*
11. *Streptococcus pneumoniae*
12. *Streptococcus haemolyticus*

Figura 13.



1. *Escherichia coli*
2. *Neisseria gonorrhoeae*
3. *Klebsiella pneumoniae*
4. *Moraxella catarrhalis*
5. *Pseudomonas aeruginosa*
6. *Proteus vulgaris*
7. *Enterococcus faecalis*
8. *Staphylococcus epidermidis*
9. *Staphylococcus aureus*
10. *Staphylococcus xylois*
11. *Streptococcus pneumoniae*
12. *Streptococcus haemolyticus*

Figura 14.



1. *Escherichia coli*
2. *Neisseria gonorrhoeae*
3. *Klebsiella pneumoniae*
4. *Moraxella catarrhalis*
5. *Pseudomonas aeruginosa*
6. *Proteus vulgaris*
7. *Enterococcus faecalis*
8. *Staphylococcus epidermidis*
9. *Staphylococcus aureus*
10. *Staphylococcus xylois*
11. *Streptococcus pneumoniae*
12. *Streptococcus haemolyticus*

Figura 15.

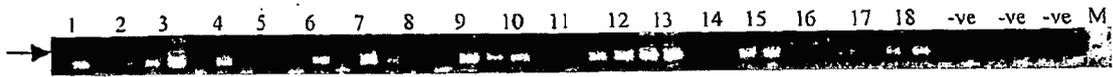


Figura 16A. *Staphylococcus epidermidis*

## Secuencia ADN normal (SEQ ID NO 223)

GATTAAGTTATTAAGGGCGCACGGTGGATGCCTTGGCACTAGAAAGCCGATGAAGGACGTTACTAACGA  
 CGATATGCTTTGGGTAGCTGTAAGTAAGCGTTGATCCAGAGATTTCCGAATGGGGGAACCCAGCATGA  
 GTTATGTCATGTTATCGATATGTGAATTTATAGCATGTCAGAAGGCAGACCCGGAGAAGCTGAAACATC  
 TTAGTACCCGGAGGAAGAGAAAAGAAAAATCGATTCCCTGAGTAGCGGCGAGCGAAAACGGGAAGAGCCC  
 AAACCAACAAGCTTGCTTGTGGGGTTGTAGGACACTCTATACGGAGTTACAAAAGAACATGTTAGAC  
 GAATCATCTGAAAAGATGAATCAAAGAAGGTAATAATCCTGTAGTCGAAAAACATATTTCTCTCTTTGAGT  
 GGATCCTGAGTACGACCGGAGCACGTGAAATTCCTGCGGAATCTGGGAGGACCATCTCCTAAGGCTAAA  
 TACTCTCTAGTGACCGATAGTGAACCAGTACCGTGAGGGAAAGGTGAAAAGTACCCCGGAAGGGGAGT  
 GAAAGAGAAGCTTGAACCGTGTGCTTACAAGTAGTCAAGAGCCCGTTAATGGGTGATGGCGTGCCTTTT  
 GTAGAAATGAACCGGCGAGTTACGATCTGATGCAAGGTTAAGCAGCAAATGCGGAGCCGACGCGAAAGC  
 GAGTCTGAATAGGGCGTTGAGTATTTGGTTCGTAGACCCGAAACAGGTGATCTACCTTGGTCAGGTT  
 GAAGTTCAGGTAACTGAATGGAGGACCGAACCGACTTACGTTGAAAAGTGAGCGGATGAACTGAGG  
 GTAGCGGAGAAAATCCAATCGAACTTGGAGATAGCTGGTCTCTCCGAAATAGCTTTAGGGCTAGCCT  
 CAAGTGATGATTATGGAGGTAGAGCACTGTTTGGACGAGGGGCCCTCTCGGGTTACCGAATTCAGA  
 CAAACTCCGAATGCCAATTAATTTAACTTGGGAGTCAAGCATGGGTGATAAGGTCCGTGTTTCAAAG  
 GGAAACAGCCAGACCACCAAGCTAAGGTCCCAAAATATATGTTAAGTGAAAAGGATGTGGCGTTGCC  
 CAGACAACTAGGATGTTGGCTTAGAAGCAGCCATATTTAAAGAGTCCGTAATAGCTCACTAGTTCGAG  
 TGACACTGCGCCGAAAAATGTACCGGGGCTAAACATATTTACCGAAGCTGTGGATTGCTCTTTGGACAAT  
 GGTAGGAGAGCGTTCTAAGGGCGTCAAGCATGATCGCAAGGACATGTGGAGCGCTTAGAAGTGAGAA  
 TGCCGGTGTGAGTAGCGAAAAGACGGGTGAGAATCCCGTCCACCGATTGACTAAGGTTTCCAGAGGAAG  
 GCTCGTCCGCTCTGGGTTAGTCCGGTCTAAGCTGAGGCCGACAGGCGTAGGCGATGGATAACAGGTT  
 GATATTCCTGTACCACCTAGTATCGTTTTAATCGATGGGGGACCGAGTAGGATAGCGAAAGCGTGCT  
 GTTGGAGTGACGTTCCAAGCAGTAAGGCTGAGTGTAGGCAAATCCGGCACTCATAAGGCTGAGCTGT  
 GATGGGGAGAGGAAATGTTTCTCGAGTCGTTGATTTCACTGCGGAGAAAAGCCTCTAGATAGAT  
 AACAGGTGCCGTTACCGCAAACCGACACAGGTAGTCAAGATGAGAATTTCTAAGGTGAGCGAGCGAACT  
 CTCGTTAAGGAACTCGGCAAAAATGACCCCGTAACTTTCGGGAGAAAGGGGTGCTCTTTAGGGTTACGCC  
 CAGAAGAGCCGAGTGAATAGGCCCAAGCGACTGTTTATCAAAAAACAGGTCTCTGTAAACCGTAA  
 GGTGATGTATAGGGGCTGACGCTGCCCCGTGCTGGAAGGTTAAGAGGAGTGGTTAGCTTCTGCGAAG  
 CTACGAATCGAAGCCCCAGTAAACGGCGCCGTAACATAACGGTCTTAAGGTAGCGAAATTCCTTGT  
 CGGGTAAGTTCGACCCCGCACGAAAGGCGTAACGATTTGGGCACTGTCTCAACGAGAGACTCGGTGAA  
 ATCATAGTACCTGTGAAGATGCAGGTTACCCGCGACAGGACGGAAAAGACCCCGTGGAGCTTTACTGTA  
 GCCTGATATTGAAATTCGGCACAGCTTGTACAGGATAGGTAGGAGCCTTTGAAAACGTGAGCGCTAGCT  
 TACGTGGAGGCGTTGGTGGGATACTACCCTAGCTGTGTTGGCTTTCTAACCCGCACCCTTATCGTGG  
 TGGGAGACAGTGTGAGGCGGCGAGTTTACTGGGGCGGTGCGCTCTAAAAGGTAACGGAGGCGCTCA  
 AAGGTTCCCTCAGAATGTTTGGAAATCATTATAGAGTGTAAAGGCATAAGGGAGCTTGACTGCGAGA  
 CCTACAAGTCGAGCAGGTCGAAAGACGGACTTAGTGATCCGGTGGTTCCGCATGGAAGGGCCATCGC  
 TCAACGGATAAAAAGCTACCCCGGGGATAACAGGCTTATCTCCCCAAGAGTTCACATCGACGGGGAGG  
 TTTGGCACCTCGATGTCGGCTCATCGCATCCTGGGGCTGTAGTCCGTTCCCAAGGGTTGGGCTGTTCGC  
 CCATTAAGCGGTACGCGAGCTGGGTTGAGAACGTGTCGTGAGACAGTTCCGGTCCCTATCCGTCGTGGGC  
 GTAGGAAATTTGAGAGGAGCTGTCTTAGTACGAGAGGACCGGGATGGACATACCTCTGGTGTACCAG  
 TTGTCGTGCCAACGGCATAGCTGGGTAGCTATGTATGGACGGGATAAGTGTGAAAGCATCTAAGCAT  
 GAAGCCCCCTCAAGATGAGATTTCCCACTTCGGTTATAAGATCCCTCGAAGATGACGAGGTTAATA  
 GGTTCGAGGTGGAAGCGTGGTGACACGTGGAGCTGACGAATACTAATCGATCGAAGACTTAATCAA

**Figura 16B. *Staphylococcus epidermidis***

Secuencia simplificada (SEQ ID NO 224 )

GATTAAGTTATTAAGGGTGTATGGTGGATGTTTTGGTATTAGAAGTTGATGAAGGATGTTATTAATGA  
 TGATATGTTTTGGGTAGTTGTAAGTAAGTGTTGATTTAGAGATTTTGAATGGGGAAATTTAGTATGA  
 GTTATGTTATGTTATGATATGTGAATTTATAGTATGTTAGAAGGTAGATTTGGAGAATGAAATATT  
 TTAGTATTTGGAGGAAGAGAAAAGAAAATGATTTTTGAGTAGTGGTGAGTGAAATGGGAAGAGTTT  
 AAATTAATAAGTTTTGTTTTGTTGGGGTTGTAGGATATTTTATATGGAGTTATAAAAAGAATAGTTAGAT  
 GAATTAATTTGGAAAGATGAATTAAGAAGGTAATAATTTTTGTAGTTGAAAATATATTTTTTTTTTCACT  
 GGATTTGAGTATGATGGAGTATGTGAAATTTGTTGGAATTTGGGAGGATTATTTTTTAAGGTTAAA  
 TATTTTTAGTGATTGATAGTGAATTAGTATTGTGAGGGAAAGGTGAAAAGTATTTTTGGAAGGGGAGT  
 GAAAGAGAATTTGAAATTTGTGTTTTATAAGTAGTTAGAGTTTGTAAATGGGTGATGGTGTGTTTTTT  
 GTAGAATGAATTTGGTGGTATGATTTGATGTAAGGTTAAGTAGTAAATGTGGAGTTGTAGTGAAGT  
 GAGTTTGAATAGGGTGTGAGTATTTGGTTGTAGATTTGAAATTAGGTGATTTATTTTTGGTTAGGTT  
 GAAGTTTACGTAATATTGAATGGAGGATTGAATTCATTTATGTTGAAAAGTGAAGTGGATGAATGAGG  
 GTAGTGGAGAAAATTTAATTTGAATTTGGAGATAGTTGGTTTTTTTTGAAAATAGTTTTAGGGTTAGTTT  
 TAAGTGATGATTATTGGAGGTAGAGTATTGTTTGGATGAGGGGTTTTTTTTGGGTTATTGAATTTAGA  
 TAAATTTGAATGTTAATTAATTTAATTTGGGAGTTAGAATATGGGTGATAAGGTTTGTGTTGAAAG  
 GCAAATAGTTTAGATTATTAGTTAAGGTTTTAAAATATATGTTAAGTGGAAAAGGATGTGGTGTGTT  
 TAGATAATTAGGATGTTGGTTTAGAAGTAGTTATTATTTAAAGAGTGTGTAATAGTTTATTAGTTGAG  
 TGATATTGTGTTGAAAATGTATTGGGGTTAAATATATTATTGAAGTTGTGGATTGTTTTTTGGATAAT  
 GGTAGGAGAGTCTTTTAAGCGTGTGAAAGTATGATTGTAAGGATATGTGGAGTGTTTAGAAGTGAGAA  
 TGTGGTGTGAGTAGTGAAGATGGGTGAGAATTTGTTTTATTGATTGATTAAGGTTTTTAGAGGAAG  
 GTTTGTTGTTTTGGGTTAGTTGGGTTTTAAGTTGAGGTTGATAGGTGATGGTATGATAATAGGTT  
 GATATTTTGTATTATTTAGTATTGTTTTAATTCATGGGGGATGACTAGGATAGGTTGAGTGTGTT  
 GTTGGAGTGTATGTTAAGTAGTAAGGTTGAGTGTAGGTAATTTGGTATTTATAAGGTTGAGTTGT  
 GATGGGGAGAGCAAATGTTTTTTTTGAGTTGTTGATTTTATATTGTTGAGAAAAGTTTTTAGATAGAT  
 AATAGGTTCTTCTATTGTAATTCATATAGCTAGTTAAGATCAGAATTTTAAGGTCAGTGAAGTGAAT  
 TTTGTTAAGGAATTTGGTAAAATGATTTTTGTAATTTTGGGAGAAGGGGTGTTTTTTAGGGTTTATGTT  
 TAGAAGAGTTGTAGTGAATAGGTTTAAAGTGAATGTTTATTAAAAATATAGGTTTTTTGTTAAATTTGTA  
 GGTGATGTATAGGGGTTGATGTTGTTTGGTGTGGAAGGTTAAGAGGAGTGGTTAGTTTTTGTGAAG  
 TTATGAATTTGAAGTTTTAGTAAATGGTGGTTGTAATTTAATGGTTTTAAGGTAGTGAATTTTTTGT  
 TGGGTAAGTTTTGATTTGTATGAAAGGTGTAATGATTTGGGTATTGTTTTAATGAGAGATTTGGTGAA  
 ATTATAGTATTTGTGAAGATGAGGTTATTTGTGATAGGATGGAAGATTTTGTGGAGTTTATTGTA  
 GTTTGATATTGAAATTTGGTATAGTTTGTATAGGATAGGTAGGAGTTTTTGAATGTGAGTGTTAGTT  
 TATGTGGAGGTGTTGGTGGGATATTATTTTAGTTGTGTTGGTTTTTTAATTTGTATTATTTATTGTGG  
 TGGGAGATAGTGTTAGGTGGGTAGTTTGAATGGGGTGGTTGTTTTTTAAAAGGTAATGGAGGTGTTTA  
 AAGGTTTTTTTAGAATGGTTGGAATTTATTTATAGAGTGTAAAGGTATAAGGGAGTTTGATTTGTGAGA  
 TTTATAAGTTGAGTAGGTTGAAAGATGGATTTAGTGATTTGGTGGTTTTGTATGGAAGGGTTATTGT  
 TTAATGGATAAAAAGTTATTTTGGGGATAATAGGTTTATTTTTTTTAAAGAGTTTATATTGATGGGGAGG  
 TTTGGTATTTTTCATGTTGGTTTATTGTATTTTGGGGTTGTAGTTGGTTTTAAGGGTTGGTTGTTTGT  
 TTATTAAGTGGTATGTGAGTTGGGTTTAGAATGTTGTGAGATAGTTTGGTTTTTTATTGTTGTGGGT  
 GTAGGAAATTTGAGAGGAGTTGTTTTTAGTATGAGAGGATGGGATGGATATATTTTTTGGTGTATTAG  
 TTGTTGTGTTAATGGTATAGTTGGGTAGTTATGTATGGATGGGATAAGTGTGAAAGTATTTAAGTAT  
 GAAGTTTTTTTTAAGATGACATTTTTTAATTTTGGTTATAAGATTTTTTGAAGATGATGAGCTTAATA  
 GTTTTGAGGTGGAAGTGTGGTATATGTGGAGTTGATCAATATTAATTTGATTGAAGATTTAATTA

**Figura 17A. Gen recA de *E. coli***

Secuencia Normal (SEQ ID NO 225)

ATGGCTATCGACGAAAAACAAACAGAAAGCGTTGGCGGCAGCACTGGGCCAGATTGAGAAACAATTTGG  
TAAAGGCTCCATCATGCGCCTGGGTGAAGACCGTTCCATGGATGTGGAAACCATCTCTACCGGTTCCG  
TTTCACTGGATATCGCGCTTGGGGCAGGTGGTCTGCCGATGGGCCGATCGTTCGAAATCTACGGACCG  
GAATCTTCCGGTAAAAACACGCTGACGCTGCAGGTGATCGCCGAGCGCAGCGTGAAGGTAAAACTG  
TGCGTTTATCGATGCTGAACACGCGCTGGACCCAATCTACGCACGTAAACTGGGCGTCGATATCGATA  
ACCTGCTGTGCTCCAGCCGGACACCGCGAGCAGGCCTGGAAATCTGTGACGCCCTGGCGCGTTCT  
GGCGCAGTAGACGTTATCGTTCGTTGACTCCGTGGCGGCCTGACGCCGAAAGCCGAAATCGAAGGCCGA  
AATCGGCGACTCTCACATGGGCCTTGGCGCACGTATGATGAGCCAGGCGATGCGTAAGCTGGCGGGTA  
ACCTGAAGCAGTCCAACACGCTGCTGATCTTCATCAACCAGATCCGTATGAAAATGGTGTGATGTTT  
GGTAAACCCGAAACCACTACCGGTGGTAAACGCGCTGAAATCTACGCCTCTGTTGCTCTCGACATCCG  
TCGTATCGGCGCGGTGAAAGAGGGCCAAAACGTGGTGGGTAGCGAAACCCGCGTAAAGTGGTGAAGA  
ACAAAATCGCTGCGCCGTTTAAACAGGCTGAATCCAGATCCTCTACGGCGAAGGTATCAACTTCTAC  
GGCGAACTGGTTGACCTGGGCGTAAAAGAGAAGCTGATCGAGAAAGCAGGCGCGTGGTACAGCTACAA  
AGGTGAGAAGATCGGTCAGGGTAAAGCGAATGCGACTGCCCTGGCTGAAAGATAACCCGAAACCCGCGA  
AAGAGATCGAGAAGAAAGTACGTGAGTTGCTGCTGAGCAACCCGAACTCAACGCCGGATTTCTCTGTA  
GATGATAGCGAAGGCTAGCAGAACTAACGAAGATTTTTAA

**Figura 17B. Gen recA de *E. coli***

Secuencia simplificada (SEQ ID NO 226)

ATGGTTATTGATGAAAATAAATAGAACTGTTGGTGGTAGTATTGGGTTAGATTGAGAAAATAATTTGG  
TAAAGGTTTTATTATGTGTTGGGTGAAGATTGTTTTATGGATGTGGAAATTAATTTTTATTGGTTTGT  
TTTTATTGGATATTGTGTTGGGTAGGTGGTTGTTGATGGGTGATTTGTTGAAATTTATGGATTG  
GAATTTTTTGGTAAAAATATGTTGATGTTGTAGGTGATTGTTGTAGTGTAGTGTGAAGGTAAAATTTG  
TGTGTTTTATTGATGTTGAATATGTGTTGGATTTAATTTATGATGTAATTTGGGTGTTGATATTGATA  
ATTTGTTGTGTTTTAGTTGGATATTGGTGAGTAGGTATTGGAATTTGTGATGTTTTGGTGTGTTTT  
CGTGTAGTAGATGTTATTGTTGTTCAATTTGTCGCTGGTATTGATGTTGAAAGTGGAAATGAAGGTGA  
AATTGGTGATTTTTATATGGGTTTTGTGGTATGTATGATGAGTTAGGTGATGTGTAAGTTGGTGGGTA  
ATTTGAAGTAGTTAATATGTTGTTGATTTTTATTAATTAGATTTGTATGAAAATTTGGTGTGATGTTT  
GGTAATTTGGAAATTAATTTGGTGGTAATGTGTTGAAATTTATGTTTTGTTGTTTTGATATTTG  
TTGTATTGGTGTGGTAAAAGAGGGTAAAATGTGGTGGTAGTAAAATTTGTGTGAAAGTGGTGAAGA  
ATAAAATTTGTTGTTGTTTAAATAGGTTGAATTTTAGATTTTTTATGGTGAAGGTATTAATTTTTAT  
GGTGAATTTGTTGATTTGGGTGAAAAGAGAAGTTGATTGAGAAAGTAGGTGTGTTGTTATAGTTATAA  
AGGTGAGAAGATTGTTAGGGTAAAGTGAATGTGATTTGTTGGTTGAAAGATAATTTGGAAATTTGTA  
AAGAGATTGAGAAGAAAGTATGTGAGTTGTTGTTGAGTAAATTTGAATTTAATGTTGATTTTTTTGTA  
GATGATAGTGAAGGTGATGAGAAATTAATGAAGATTTTTAA