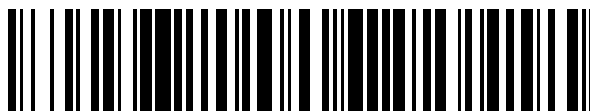


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 587 595**

51 Int. Cl.:

C07K 7/08 (2006.01)

C07K 14/00 (2006.01)

C12N 15/11 (2006.01)

C07K 7/06 (2006.01)

C07K 19/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **25.09.2008 PCT/NZ2008/000247**

87 Fecha y número de publicación internacional: **02.04.2009 WO09041830**

96 Fecha de presentación y número de la solicitud europea: **25.09.2008 E 08834203 (5)**

97 Fecha y número de publicación de la concesión europea: **06.04.2016 EP 2203466**

54 Título: **Péptidos y polipéptidos de permeabilización celular para células microbianas**

30 Prioridad:

25.09.2007 US 975104 P

22.11.2007 US 989840 P

22.11.2007 US 989841 P

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

25.10.2016

73 Titular/es:

**PASTORAL GREENHOUSE GAS RESEARCH LTD
(100.0%)**

**Level 4, Wellington Chambers, 154 Featherston
Street**

Wellington 6140, NZ

72 Inventor/es:

ATTWOOD, GRAEME TREVOR;

KELLY, WILLIAM JOHN y

ALTERMANN, ERIC HEINZ

74 Agente/Representante:

ZEA CHECA, Bernabé

ES 2 587 595 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

Péptidos y polipéptidos de permeabilización celular para células microbianas

5 Campo de la Invención

La invención se relaciona con composiciones y métodos para aplicar moléculas inhibitorias al interior de las células microbianas, especialmente de células de metanógenos. Específicamente, la invención se relaciona con péptidos señal y con los polipéptidos que contienen estos péptidos, así como con polinucleótidos que codifican estos péptidos o polipéptidos. La invención se relaciona asimismo con vectores de expresión y células huésped para la producción de estos péptidos o polipéptido. La invención se relaciona también con métodos para la detección, direccionamiento, permeabilización e inhibición de células microbianas, especialmente células metanógenas, utilizando los péptidos o polipéptidos, vectores de expresión y células huésped descritos.

15 Antecedentes de la Invención

En Nueva Zelanda, la actividad agrícola es responsable de la mayor parte de las emisiones de gas de invernadero. Por lo tanto, la reducción de las emisiones agrícolas de gases invernadero es importante para que Nueva Zelanda cumpla las obligaciones surgidas del Protocolo de Kioto. El Protocolo requiere la reducción de los gases invernadero a los niveles de 1990 antes del final del primer período de compromiso (2008–2012). Para este fin, algunos grupos del sector agrícola y el gobierno de Nueva Zelanda establecieron el Consorcio Pastoral de Investigaciones sobre Gases invernadero (PGGRC) para identificar medios para reducir las emisiones de gases invernadero en la agricultura de Nueva Zelanda.

Una parte importante de las actividades del PGGRC ha consistido en investigar la forma de reducir las emisiones de metano producidas por los rumiantes que pastan en Nueva Zelanda. La mitigación de las emisiones de metano de los rumiantes es de interés comercial por dos razones. En primer lugar, la falta de cumplimiento de los compromisos asumidos bajo el Protocolo de Kioto obligará al gobierno a adquirir créditos de carbono. Se estima que actualmente el costo es de \$350 millones. En segundo lugar, la producción de metano da lugar a la pérdida de 8–12% de la energía bruta producida en el rumen. Esta energía se podría utilizar, en su lugar, para mejorar la productividad de los rumiantes.

El metano se produce en el rumen merced a los microbios denominados metanógenos que son parte del filum Euryarchaeota dentro del reino de los Archaea. La mayoría de los metanógenos se desarrollan en CO₂ y H₂ como única fuente de energía, aunque pueden utilizar compuestos de acetato o metilo para desarrollarse. En el rumen existen varios géneros diferentes de archaea metanogénicos, aunque se piensa que algunas especies del género *Methanobrevibacter*, especialmente *M. ruminantium* y *M. smithii* son los metanógenos predominantes en los rumiantes de Nueva Zelanda. En la actualidad, *M. ruminantium* es el tema central de un proyecto de secuenciación de genoma con fondos otorgados por el PGGRC. El proyecto es la primera secuenciación del genoma de un metanógeno del rumen y apunta a cimentar una mejor comprensión de la biología de *Methanobrevibacter* para descubrir blancos para la inhibición de la formación de metano.

La reducción de la producción de metano en el rumen requiere la inhibición de los metanógenos o la inactivación de su vía de metanogénesis. Un medio para inhibir la producción de metano consiste en aplicar moléculas inhibitorias específicas a las células metanógenas. Esto se puede lograr, por ejemplo, acoplando moléculas inhibitorias a péptidos impermeabilizantes de las células. En las células microbianas, los péptidos señal median en la translocación de las proteínas extracelulares desde el interior al exterior de la célula y son adecuados para el transporte de las moléculas inhibitorias. Por lo tanto, sería provechoso identificar péptidos señal que tengan la capacidad de permeabilizar las células metanógenas y aplicar inhibidores.

Los péptidos señal, o secuencias señal, están incluidos típicamente en las proteínas precursoras secretadas por las células procariotas y eucariotas. Los péptidos señal son parte de una extensión permeabilizante de las células que existe en el término N del precursor. La secuencia primaria de aminoácidos de los péptidos señal no se conserva aparte del sitio de escisión de la peptidasa señal (von Heijne, 1985). Sin embargo, los péptidos señal sí comparten similitudes estructurales. Por lo general, los péptidos señal incluyen de uno a cinco residuos aminoácidos N-terminales con carga positiva (región n) seguidos por 10 a 15 residuos aminoácidos hidrófobos (región h).

Habitualmente hay un residuo de glicina o prolina situado dentro del dominio hidrófobo y uno o más residuos de treonina y/o serina forman un dominio polar (región c) cerca del sitio de escisión (Inouye y Haleboua, 1980; Vlasuk et al., 1983, von Heijne, 1985).

Se ha propuesto un modelo de bucle para la translocación de los péptidos señal (Inouye et al., 1977; Inouye y Halagoua, 1980) por el cual el término N con carga positiva del péptido señal interactúa con la superficie interna con

carga negativa de la membrana celular. De esa manera, el dominio hidrófobo es atraído hacia la bicapa lipídica hidrófoba de la membrana, formando un bucle. Eventualmente el bucle incluye el sitio de escisión, que queda expuesto a la peptidasa señal para la eliminación del péptido señal. Una de las barreras para la inhibición o limitación de la formación de metano es la capacidad de aplicar compuestos inhibitorios al interior de las células metanógenas. Por consiguiente, existe la necesidad de identificar péptidos señal que puedan adherirse a las membranas celulares y transportar moléculas a través de la bicapa lipídica, como portadores útiles de inhibidores celulares.

Breve Descripción de la Invención

10

La invención presenta un péptido señal o un polipéptido que comprende este péptido, que comprende por lo menos una secuencia de aminoácidos seleccionada del grupo que consiste en SEC. ID. NO: 1-172. En un aspecto específico, el péptido o polipéptido comprende por lo menos una secuencia de aminoácidos KKLIIIIIIIIIIIIIIIIII de SEC. ID. NO: 117, o por lo menos una secuencia de aminoácidos KKIIIIIIIIIIIIIIIIII de SEC. ID. NO: 119. En otro aspecto, el péptido o polipéptido comprende un fragmento que incluye, por ejemplo, por lo menos una secuencia de aminoácidos que comprende los aminoácidos 3-14, 3-16 o 2-16 de SEC. ID. NO: 117, o por lo menos una secuencia de aminoácidos que comprende los aminoácidos 3-15, 3-17 o 2-17 de SEC. ID. NO: 119. En otro aspecto, el péptido o polipéptido comprende un fragmento que incluye por lo menos una secuencia núcleo conservada de acuerdo con SEC. ID. NO: 1-172, como se describe en la presente. En un aspecto más, el péptido o polipéptido es codificado por al menos un fragmento de un polinucleótido seleccionado del grupo que consiste en SEC. ID. NO: 173-341 o SEC. ID. NO: 342-533.

15

20

La invención presenta asimismo un polinucleótido aislado que comprende una secuencia codificadora para por lo menos un péptido señal o un polipéptido que comprende este péptido. En un aspecto, el polinucleótido comprende una secuencia codificadora para por lo menos una secuencia de aminoácidos seleccionada del grupo que consiste en SEC. ID. NO: 1-172. En un aspecto específico, el polinucleótido comprende una secuencia codificadora para por lo menos una secuencia de aminoácidos KKLIIIIIIIIIIIIIIIIII de SEC. ID. NO: 117, o una secuencia codificadora para al menos una secuencia de aminoácidos KKIIIIIIIIIIIIIIIIII de SEC. ID. NO: 119. En otro aspecto, el polinucleótido comprende un fragmento de una secuencia codificadora, por ejemplo una secuencia codificadora para por lo menos una secuencia de aminoácidos que comprende los aminoácidos 3-14, 3-16 o 2-16 de SEC. ID. NO: 117, o una secuencia codificadora, por ejemplo una secuencia codificadora para por lo menos una secuencia de aminoácidos que comprende los aminoácidos 3-15, 3-17 o 2-17 de SEC. ID. NO: 119. En un aspecto adicional, el polinucleótido comprende un fragmento de una secuencia codificadora, por ejemplo una secuencia nucleotídica que codifica por lo menos una secuencia núcleo conservada correspondiente a SEC. ID. NO: 1-172, de acuerdo con lo descrito en la presente.

25

30

35

En un aspecto adicional, la invención presenta un polinucleótido aislado que comprende una secuencia de ácidos nucleicos seleccionada del grupo que consiste en SEC. ID. NO: 173-341 o SEC. ID. NO: 342-533. En un aspecto específico, el polinucleótido comprende la secuencia de ácidos nucleicos de SEC. ID. NO: 531, 532, o 533. En otro aspecto, el polinucleótido es un fragmento o un oligonucleótido que comprende, por ejemplo, la secuencia de ácidos nucleicos que se extiende desde el nucleótido 7-42, 7-48 o 4-48 de SEC. ID. NO: 531, 532 o 533. Además, la invención abarca un polinucleótido aislado, o un fragmento del mismo, que se hibrida a cualquiera de las secuencias de ácidos nucleicos de SEC. ID. NO: 173-341 o SEC. ID. NO: 342-533. La invención abarca además un polinucleótido aislado que comprende el complemento, complemento inverso, secuencia inversa o fragmentos de la misma, de cualquiera de las secuencias de ácidos nucleicos que codifican un péptido señal ó un polipéptido que comprende este péptido.

40

45

La invención presenta un vector de expresión que comprende un polinucleótido que comprende una secuencia codificadora para al menos un péptido señal o un polipéptido que comprende este péptido. En un aspecto, el vector de expresión comprende una secuencia codificadora para por lo menos una secuencia de aminoácidos seleccionada del grupo que consiste en SEC. ID. NO: 1-172. En un aspecto específico, el vector de expresión comprende una secuencia codificadora para por lo menos una secuencia de aminoácidos KKLIIIIIIIIIIIIIIIIII de SEC. ID. NO: 117, o una secuencia codificadora para por lo menos una secuencia de aminoácidos KKIIIIIIIIIIIIIIIIII de SEC. ID. NO: 119. En otro aspecto, el vector de expresión comprende una secuencia codificadora para por lo menos una secuencia de aminoácidos que se extiende desde el aminoácido 3-14, 3-16 o 2-16 de SEC. ID. NO: 117 o una secuencia codificadora para por lo menos una secuencia de aminoácidos que comprende los aminoácidos 3-15, 3-17 o 2-17 de SEC. ID. NO: 119. En un aspecto más, la invención presenta una célula huésped, por ejemplo una célula huésped microbiana, que comprende por lo menos un vector de expresión.

50

55

La invención presenta específicamente un anticuerpo dirigido a un péptido, polipéptido o polinucleótido descrito en este documento. En ciertos aspectos, el anticuerpo se dirige a por lo menos una secuencia de péptido señal seleccionada del grupo que consiste en SEC. ID. NO: 1-172 o una secuencia modificada de la misma. En otros aspectos, el anticuerpo se dirige a por lo menos un fragmento de una secuencia peptídica señal, por ejemplo una

60

- secuencia núcleo conservada de una secuencia seleccionada del grupo que consiste en SEC. ID. NO: 1–172. En otro aspecto, el anticuerpo se une a un polipéptido que comprende una secuencia peptídica señal de acuerdo con cualquiera de SEC. ID. NO: 1–172. En otros aspectos, el anticuerpo se dirige a por lo menos un fragmento de un polinucleótido seleccionado del grupo que consiste en SEC. ID. NO: 173–341 o SEC. ID. NO: 342–533 o un complemento o secuencia modificada de los mismos. En otro aspecto, el anticuerpo incluye una o más fusiones o conjugados con por lo menos un inhibidor celular, por ejemplo compuestos antimetanolgólisis (por ejemplo ácido bromoetansulfónico), anticuerpos y fragmentos de anticuerpos, enzimas líticas, ácidos nucleicos peptídicos, péptidos antimicrobianos y otros antibióticos, como se describe en la presente en forma detallada.
- 5
- 10 La invención presenta además péptidos señal modificados y polipéptidos que comprenden estos péptidos, así como anticuerpos dirigidos a estos péptidos o polipéptidos, incluyendo las alternaciones, fragmentos, variantes y derivados biológicamente activos descritos en el presente documento. También se presentan polinucleótidos que codifican estos péptidos o polipéptidos modificados, como así también alteraciones, fragmentos, variantes y derivados de los polinucleótidos descritos, vectores de expresión que comprenden estas secuencias de ácido nucleico y células huésped que comprenden estos vectores. En aspectos específicos, las composiciones y métodos de la presente invención emplean estos polinucleótidos, polipéptidos o anticuerpos modificados o los correspondientes vectores de expresión o células huésped. En aspectos específicos, los péptidos o polipéptidos se producen en forma de fusiones o conjugados con por lo menos un inhibidor celular, por ejemplo compuestos antimetanolgólisis (por ejemplo ácido bromoetansulfónico), anticuerpos y fragmentos de anticuerpos, enzimas líticas, ácidos nucleicos peptídicos, péptidos antimicrobianos y otros antibióticos de acuerdo con lo descrito en forma detallada en la presente.
- 15
- 20

La invención presenta además una composición que comprende un péptido señal aislado (por ejemplo, por lo menos una de SEC. ID. NO: 1–172 o una secuencia modificada de las mismas) o un polipéptido que comprende este péptido o un anticuerpo dirigido a este péptido o polipéptido. También se presenta una composición que comprende un polinucleótido aislado (por ejemplo, por lo menos uno de SEC. ID. NO: 173–341 o SEC. ID. NO: 342–533 o un complemento o secuencia modificada de las mismas). Se presenta asimismo una composición que incluye un vector de expresión o una célula huésped que comprende un vector de expresión de acuerdo con la presente invención. La composición puede incluir cualquiera de las alteraciones, fragmentos, variantes biológicamente activas y derivados descritos en la presente. Las composiciones pueden comprender además por lo menos un inhibidor celular y pueden ser formuladas, por ejemplo, en forma de composiciones farmacéuticas o como suplementos alimentarios, específicamente componentes del alimento para rumiantes.

25

30

En un aspecto específico, la invención presenta una composición de la presente invención como parte de un kit para la detección y/o medición o direccionamiento, permeabilización y/o inhibición de células microbianas, especialmente células metanógenas, de acuerdo con los métodos descritos. Los kits comprenden: a) por lo menos una composición descrita en este documento y b) optativamente, instrucciones para su uso, por ejemplo en el direccionamiento o permeabilización de las células o la inhibición del desarrollo o replicación de las células para los metanógenos u otros microbios. En aspectos específicos, el péptido o polipéptido comprende por lo menos una secuencia de aminoácidos seleccionada del grupo que consiste en SEC. ID. NO: 1–172 o una secuencia modificada de las mismas.

35

40

La invención presenta un método para producir un péptido señal o un polipéptido que comprende este péptido, método que comprende: a) cultivar un vector de expresión o célula huésped que comprende un vector de expresión, que comprende una secuencia codificadora para por lo menos un péptido señal o un polipéptido que comprende este péptido, en condiciones adecuadas para la expresión del péptido o polipéptido y b) recuperar el péptido o polipéptido del cultivo. También se presentan métodos para la producción de las composiciones descritas. En otros aspectos, el péptido o polipéptido comprende por lo menos una secuencia de aminoácidos seleccionada del grupo que consiste en SEC. ID. NO: 1–172 o una secuencia modificada de las mismas.

45

La invención presenta además un método para producir un péptido señal o un polipéptido que comprende este péptido, que incluye una fusión o conjugado con por lo menos un inhibidor celular, por ejemplo compuestos antimetanolgólisis (por ejemplo ácido bromoetansulfónico), anticuerpos y fragmentos de anticuerpos, enzimas líticas, ácidos nucleicos peptídicos, péptidos antimicrobianos y otros antibióticos, como se describe en la presente en forma detallada. Dicho método comprende: a) cultivar un vector de expresión o una célula huésped que comprende un vector de expresión, que comprende una secuencia codificadora para por lo menos un péptido o polipéptido en condiciones adecuadas para la expresión del péptido o polipéptido, b) formar la fusión o conjugado (por ejemplo, mediante la expresión de la secuencia fusionada o por conjugación química al inhibidor celular) y c) recuperar la fusión o conjugado. En aspectos específicos, el péptido señal o polipéptido comprende por lo menos una secuencia de aminoácidos seleccionada del grupo que consiste en SEC. ID. NO: 1–172 o una secuencia modificada de las mismas.

50

55

60

La invención presenta un método para permeabilizar una célula microbiana, especialmente una célula metanógena, que comprende: a) optativamente, producir o aislar por lo menos un péptido señal o un polipéptido que comprende

este péptido y b) poner en contacto a la célula con el péptido señal o polipéptido. En un aspecto específico, el péptido o polipéptido comprende por lo menos una secuencia de aminoácidos seleccionada del grupo que consiste en SEC. ID. NO: 1–172 o una secuencia modificada de las mismas. En otros aspectos, el péptido o polipéptido comprende una fusión o conjugado con por lo menos un inhibidor celular, por ejemplo compuestos
5 antimetanogénesis (por ejemplo ácido bromoetansulfónico), anticuerpos y fragmentos de anticuerpos, enzimas líticas, ácidos nucleicos peptídicos, péptidos antimicrobianos y otros antibióticos, como se describe en la presente en forma detallada.

La invención presenta además un método para inhibir una célula microbiana (por ejemplo, inhibir el desarrollo o
10 replicación), específicamente una célula metanógena, que comprende: a) optativamente, producir o aislar por lo menos un péptido señal o un polipéptido que comprende este péptido, que a su vez comprende por lo menos un inhibidor celular y b) poner en contacto a la célula con el péptido señal o polipéptido. En un aspecto específico, el péptido o polipéptido comprende por lo menos una secuencia de aminoácidos seleccionada del grupo que consiste en SEC. ID. NO: 1–172 o una secuencia modificada de las mismas. En otro aspecto, el inhibidor celular es
15 seleccionado entre compuestos antimetanogénesis (por ejemplo ácido bromoetansulfónico), anticuerpos y fragmentos de anticuerpos, enzimas líticas, ácidos nucleicos peptídicos, péptidos antimicrobianos y otros antibióticos, como se describe en la presente en forma detallada.

La invención presenta asimismo un método para inhibir una célula microbiana (por ejemplo, inhibir el desarrollo o
20 replicación), específicamente una célula metanógena, que comprende: a) optativamente, producir o aislar por lo menos un péptido señal o un polipéptido que comprende este péptido, que a su vez comprende por lo menos un inhibidor celular y b) poner en contacto a la célula con el péptido señal o polipéptido. En un aspecto específico, el péptido o polipéptido comprende por lo menos una secuencia de aminoácidos seleccionada del grupo que consiste en SEC. ID. NO: 1–172 o una secuencia modificada de las mismas. En otro aspecto, el inhibidor celular es
25 seleccionado entre compuestos antimetanogénesis (por ejemplo ácido bromoetansulfónico), anticuerpos y fragmentos de anticuerpos, enzimas líticas, ácidos nucleicos peptídicos, péptidos antimicrobianos y otros antibióticos, como se describe en la presente en forma detallada.

La invención presenta también un método para la detección y/o medición de los niveles de un péptido señal o un
30 polipéptido o polinucleótido correspondiente, que comprende: 1) poner en contacto a una muestra obtenida de un sujeto con un anticuerpo dirigido a un péptido señal (por ejemplo, por lo menos uno de SEC. ID. NO: 1–172 o una secuencia modificada de las mismas) o un polipéptido o polinucleótido correspondiente y 2) determinar la presencia o niveles del complejo anticuerpo formado con el péptido señal o un polipéptido o polinucleótido correspondiente en la muestra. Dichos métodos se pueden utilizar asimismo para la detección y/o medición de los niveles de una célula
35 microbiana, especialmente una célula metanógena.

La invención presenta además un método para la detección y/o medición de los niveles de una secuencia de polinucleótidos señal (por ejemplo, una secuencia codificadora de un péptido señal o una secuencia codificadora de un polipéptido correspondiente), que comprende: 1) poner en contacto a una muestra obtenida de un sujeto con un
40 polinucleótido complementario (por ejemplo una secuencia complementaria de cualquiera de SEC. ID. NO: 173–341 o una secuencia modificada de las mismas) y 2) determinar la presencia o niveles del complejo de hibridación formado con el polinucleótido de la secuencia señal en la muestra. Dichos métodos se pueden utilizar asimismo para la detección y/o medición de los niveles de una célula microbiana, especialmente una célula metanógena.

45 En aspectos específicos, los métodos de la presente invención utilizan componentes de expresión in vivo o in vitro. En otros aspectos, los métodos emplean péptidos o polipéptidos producidos por medios o péptidos o polipéptidos recombinantes, sintéticos o semisintéticos producidos por medios endógenos.

A continuación se describen otros aspectos y modalidades de la presente invención.

50 Breve Descripción de las Figuras

Se describe la presente invención con referencia a modalidades específicas de la misma y con referencia a las Figuras.

55 Figuras 1A–1C: Comparación de genomas metanobacterianos (Figura 1A); estadísticas correspondientes al genoma de *M. ruminantium* (Figura 1B); Genes que se consideran implicados en la metanogénesis en las especies Metanobacterianas (Figura 1C).

60 Figura 2: Alineación del péptido señal de *Methanobrevibacter ruminantium*. La región núcleo conservada de cada péptido se presenta en negrita.

Figura. 3A: Logo de la secuencia proteica de 102 secuencias generadas empleando LogoBar. Figura 3B: Logo de la secuencia proteica de 102 secuencias generadas empleando LogoBar, que ilustra los residuos aminoácidos más conservados. Figura 3C: Secuencia del péptido señal consensual núcleo correspondiente a *M. ruminantium*. Figura 3D: secuencia de aminoácidos de un péptido permeabilizador de células de *M. ruminantium* con la adición de lisina-5 fluoresceína N-terminal.

Figura 4: Permeabilización de células de *M. ruminantium* con un péptido marcado con fluoresceína.

Figura 5: Diagrama de Venn que ilustra las predicciones de péptidos señal de SignalP 3.0-HMM usando tres modelos diferentes correspondientes al orfeoma señal M1093 de *M. ruminantium*.

Figura 6: Genes de *M. ruminantium* y los correspondientes puntajes de péptido señal.

Figura 7: Genes de *M. ruminantium* y los correspondientes péptidos señal.

Figura 8: Secuencias codificadoras para los péptidos señal de la FIG. 7.

Figura 9: Secuencias codificadoras para los péptidos señal de la figura 7 con codones optimizados para la expresión en *E. coli*.

Descripción Detallada de la Invención

Definiciones

Las secuencias de ácido nucleico "alteradas" que codifican péptidos señal incluyen, en el presente contexto, las que tienen delecciones, inserciones o sustituciones de diferentes nucleótidos, lo que da origen a un polinucleótido que codifica péptidos iguales o funcionalmente equivalentes. El péptido codificado también puede ser "alterado" y contener delecciones, inserciones o sustituciones de residuos aminoácidos, lo que produce un cambio silencioso y de lugar a un péptido funcionalmente equivalente. Se pueden efectuar sustituciones de aminoácidos deliberadas basándose en la similitud de polaridad, carga, solubilidad, hidrofobicidad, hidrofiliidad y/o la naturaleza anfipática de los residuos, siempre que se mantenga la actividad biológica (por ejemplo, la asociación a las células o la permeabilización de las células) o la actividad inmunogénica del péptido. Por ejemplo, los aminoácidos con carga negativa pueden incluir ácido aspártico y ácido glutámico; los aminoácidos con carga positiva pueden incluir lisina y arginina y los aminoácidos con grupos al frente polares sin carga con valores de hidrofiliidad similares pueden incluir leucina, isoleucina y valina, glicina y alanina, asparagina y glutamina, serina y treonina y fanilalanina y tirosina. "Secuencia de aminoácidos", se refiere, en el presente contexto, a la secuencia de un oligopéptido, péptido, polipéptido o proteína, y fragmentos de los mismos, y a moléculas de origen natural, sintético o semisintético. Las secuencias de la presente invención (por ejemplo, SEC. ID. NO: 1-172) comprenden por lo menos 5, 6, 7, 8, 9, 10, 11, 12, 15, 17, 19 o 22 aminoácidos, preferentemente por lo menos 5 a 10, 5 a 15, 10 a 15, 12 a 15, 15 a 17, 17 a 19 o 17 a 22 aminoácidos y que, preferentemente, retienen la actividad biológica (por ejemplo, la asociación a las células o la permeabilización de las células) o la actividad inmunológica (por ejemplo un sitio de unión al anticuerpo) de la secuencia original. Cuando en la presente se menciona la frase "secuencia de aminoácidos" para referirse a una secuencia de aminoácidos de una molécula peptídica o polipeptídica de origen natural, dicha expresión secuencia de aminoácidos o términos similares no tienen por fin limitar la secuencia de aminoácidos a la secuencia de aminoácidos nativa completa asociada a la molécula en toda su extensión.

El término "amplificación", empleado en la presente se refiere a la producción de copias adicionales de una secuencia de ácido nucleico y en general se lleva a cabo utilizando tecnologías de reacción en cadena de la polimerasa (PCR) muy conocidas en la técnica (Dieffenbach, C. W. y G. S. Dveksler (1995) PCR Primer, a Laboratory Manual, Cold Spring Harbor Press, Plainview, NY).

El término "anticuerpo" debe ser interpretado en su sentido más amplio posible y ha de incluir anticuerpos monoclonales intactos y anticuerpos policlonales. También se pretende que cubra fragmentos y derivados de anticuerpos, siempre que exhiban la actividad biológica pretendida. Los anticuerpos abarcan moléculas de inmunoglobulina y porciones activas de moléculas de inmunoglobulina (Ig), es decir, moléculas que contienen un sitio de unión al antígeno que se une específicamente (inmunorreacciona con) un antígeno. Entre estas se incluyen, aunque no a modo de limitación, los policlonales, monoclonales, quiméricos, monocatenarios, fragmentos Fc, Fab, Fab', y Fab2 y una biblioteca de expresión de Fab.

La expresión moléculas de anticuerpos se relaciona con cualquiera de las clases IgG, IgM, IgA, IgE y IgD, que difieren entre sí por la naturaleza de la cadena pesada presente en la molécula. Estas incluyen también las subclases, como por ejemplo IgG1, IgG2, y otras. La cadena liviana puede ser una cadena kappa o lambda. La referencia a anticuerpos en la presente incluye la referencia a todas las clases, subclases y tipos. También se

incluyen los anticuerpos quiméricos, por ejemplo anticuerpos monoclonales o fragmentos de los mismos que son específicos para más de un origen, por ejemplo una o más secuencias de ratón, humanos o rumiantes. También se incluyen los anticuerpos o nanocuerpos de camélidos. Se ha de entender que cada referencia a "anticuerpos" o cualquier término similar, incluye anticuerpos intactos, así como cualquier fragmento, alteración, derivado o variante de los mismos.

Los términos "biológicamente activo" o "funcional" empleados en este contexto se refieren a un péptido o polipéptido que retiene una o más funciones estructurales, inmunogénicas o bioquímicas (por ejemplo, la asociación a las células o la permeabilización de las células) de una secuencia de origen natural. Por ejemplo, en un caso, una secuencia funcional puede comprender por lo menos una de las regiones núcleo conservadas descritas en la presente.

Los términos "inhibidor celular" o "inhibidor" empleados en la presente se refieren a agentes que reducen o bloquean el crecimiento o replicación de células microbianas, especialmente células metanógenas. Un inhibidor celular puede actuar reduciendo o bloqueando, por ejemplo, la división celular. Un inhibidor puede reducir o bloquear, por ejemplo, la síntesis de ADN, la síntesis de ARN, la síntesis proteica o las modificaciones post-traducción. Un inhibidor también puede reducir o bloquear la actividad de las enzimas implicadas en la vía de la metanogénesis. Un inhibidor también puede orientar a una célula para el reconocimiento por componentes del sistema inmune. La inhibición de una célula también incluye la muerte celular, por ejemplo como resultado de lisis, apoptosis, necrosis, etc. Entre los inhibidores provechosos se cuentan, aunque no a modo de limitación, los compuestos antimetanogénesis (por ejemplo ácido bromoetansulfónico), anticuerpos y fragmentos de anticuerpos, enzimas líticas, ácidos nucleicos peptídicos, péptidos antimicrobianos y otros antibióticos, como se describe en la presente en forma detallada.

Los términos "complementarios" o "complementariedad" empleados en este contexto, se refieren a la unión natural de polinucleótidos en condiciones de sal y temperatura permisibles mediante apareamiento de bases. En el caso de la secuencia A-G-T, la secuencia complementaria es T-C-A, el complemento inverso es A-C-T y la secuencia inversa es T-G-A. La complementariedad entre dos moléculas de cadena simple puede ser parcial, donde sólo parte del ácido nucleico se une o puede ser completa cuando existe complementariedad total entre las moléculas de cadena simple. El grado de complementariedad entre hebras de ácido nucleico tiene efectos significativos sobre la eficiencia y la potencia de la hibridación entre hebras de ácido nucleico. Esto es de particular importancia en las reacciones de amplificación, que dependen de la unión entre hebras de ácidos nucleicos y en el diseño y uso de moléculas de PNA.

El término "derivado", utilizado en el presente contexto, se refiere a la modificación química de un ácido nucleico que codifica un péptido señal o de un ácido nucleico complementario del mismo. Dichas modificaciones incluyen, por ejemplo, el reemplazo del hidrógeno por un grupo alquilo, acilo o amino. En los aspectos preferidos, un derivado de ácido nucleico codifica un péptido que retiene la función biológica o inmunológica de la molécula natural. Un derivado peptídico es aquel que ha sido modificado por glicosilación, pegilación o cualquier proceso similar y que retiene una o más funciones biológicas (por ejemplo, la asociación a las células y la permeabilización de las células) o función inmunogénica de la secuencia de la cual se la derivara.

El término "homología" empleado en la presente se refiere a un grado de complementariedad. Puede haber homología parcial (es decir, menos de 100 % de identidad) u homología completa (es decir, 100 % de identidad). Se hace referencia a una secuencia parcialmente complementaria que inhibe, por lo menos parcialmente, una secuencia idéntica impidiendo que se hibride a un ácido nucleico objetivo utilizando el término funcional "sustancialmente homóloga". La inhibición de la hibridación de la secuencia totalmente complementaria a la secuencia objetivo se puede examinar empleando un ensayo de hibridación (transferencia Southern o Northern, hibridación en solución y demás) en condiciones de baja rigurosidad. Una secuencia o sonda de hibridación sustancialmente homóloga compite por la unión de una secuencia completamente homóloga a la secuencia objetivo, y la inhibe, en condiciones de baja rigurosidad. Esto no significa que las condiciones de baja rigurosidad sean tales como se permita la unión no específica; las condiciones de baja rigurosidad requieren que la unión de las dos secuencias entre sí represente una interacción específica (es decir, selectiva).

El término "hibridación", utilizado en el presente contexto, se refiere a cualquier proceso por el cual una hebra de ácido nucleico se une a una hebra complementaria por medio de apareamiento de bases.

Una "inserción" o "adición", se refiere, en el presente contexto, a un cambio de la secuencia de aminoácidos o nucleótidos como resultado de la adición de uno o más residuos aminoácidos o nucleótidos, respectivamente, en comparación con la molécula de origen natural.

Un "metanógeno" se refiere, en este contexto, a microbios que producen gas metano, entre los que se incluyen Methanobrevibacter, Methanothermobacter, Methanomicrobium, Methanobacterium y Methanosarcina. Entre los metanógenos específicos se incluyen, aunque no a modo de limitación, Methanobrevibacter ruminantium,

Methanobrevibacter smithii, Methanobrevibacter acididurans, Methanobrevibacter thaueri, Methanobacterium bryantii, Methanobacterium formicicum, Methanothermobacter marburgensis, Methanothermobacter wolfeii, Methanosphaera stadtmanae, Methanomicrobium mobile, Methanosarcina barkeri, Methanosarcina mazei, Methanococcoides burtonii y Methanolobus taylorii. Todos los géneros y especies de metanógenos están abarcados por este término.

La expresión células "microbianas" empleada en la presente se refiere a células microbianas de origen natural o genéticamente modificadas que incluyen arqueobacterias tales como metanógenos, halófilos y termoacidófilos y cianobacterias, espiroquetas, proteobacterias, como así también bacterias gram positivas y gram negativas.

10 El término "modificado" se refiere a secuencias alteradas y a fragmentos, variantes y derivados de secuencias, de acuerdo con lo descrito en la presente.

La expresión "secuencia de ácido nucleico" o "secuencia nucleotídica" empleada en este contexto, se refiere a una secuencia de un polinucleótido, oligonucleótido o fragmentos de los mismos y al ADN o ARN de origen natural, recombinante, sintético o semisintético que puede ser mono- o bicatenario, y puede representar la hebra en sentido o antisentido y las regiones codificadoras y no codificadoras. Las secuencias de la presente invención incluyen, muy preferentemente, secuencias codificadoras de polipéptidos (por ejemplo, SEC. ID. NO: 173-341 o 342-533 o complementos o secuencias modificadas de las mismas) que comprenden por lo menos 15, 18, 21, 24, 27, 30, 33, 36, 39, 45, 51, 57 o 66 nucleótidos, preferentemente por lo menos 15 a 30, 15 a 45, 30 a 45, 36 a 45, 45 a 51, 51 a 57 o 51 a 66 nucleótidos o por lo menos 100 nucleótidos o por lo menos 1000 nucleótidos. Se ha de entender que cada referencia a una "secuencia de ácido nucleico" o "secuencia nucleotídica" en este contexto debe incluir la secuencia completa nativa (por ejemplo, SEC. ID. NO: 173-341 o 342-533), así como todo complemento, fragmento, alteración, derivado o variante de la misma.

25 El término "oligonucleótido" se refiere a una secuencia de ácido nucleico que comprende por lo menos 6, 8, 10, 12, 15, 18, 21, 25, 27, 30 o 36 nucleótidos o por lo menos 12 a 36 nucleótidos o por lo menos 15 a 30 nucleótidos (por ejemplo, por lo menos un fragmento de SEC. ID. NO: 173-341 o 342-533 o un complemento de la misma), que se puede utilizar en ensayos de amplificación por PCR, secuenciación o hibridación. En el presente contexto, el término oligonucleótido es sustancialmente equivalente a los términos "amplímeros", "cebadores", "oligómeros", "oligos" y "sondas" de acuerdo con las definiciones corrientes en la técnica.

El término "polinucleótido," utilizado en singular o plural, se refiere en general a cualquier secuencia de ácido nucleico, por ejemplo cualquier polirribonucleótido o polidesoxribonucleótido, que puede ser ARN o ADN no modificado o ARN o ADN modificado. Esto incluye, aunque no a modo de limitación, ADN mono y bicatenario, ADN que incluye regiones mono y bicatenarias, moléculas híbridas que comprenden ADN y ARN que pueden ser monocatenarios o, más típicamente, bicatenarios o incluir regiones mono y bicatenarias. También se incluyen regiones de cadena triple que comprenden ARN o ADN o tanto ADN como ARN. Se incluyen específicamente ARNm, ADNc y ADN genómicos y cualquier fragmento de los mismos. El término incluye ADN y ARN que contienen una o más bases modificadas, tales como bases tritadas o bases inusuales tales como inosina. Los polinucleótidos de la presente invención pueden abarcar secuencias codificadoras o no codificadoras o secuencias en sentido o antisentido o ARNi tales como ARNsi. Se ha de entender que cada referencia a un "polinucleótido" o término similar incluye las secuencias de extensión total así como cualquier complemento, fragmento, alteración, derivado o variante de las mismas.

45 "Péptido ácido nucleico" o "PNA" utilizado en la presente se refiere a una molécula antisentido o agente antigénico que comprende bases ligadas por medio de un esqueleto peptídico.

El término "rumiante," empleado en este contexto se refiere a animales que tienen un rumen como tipo especial de órgano digestivo. Entre los rumiantes se incluyen, aunque no a modo de limitación, vacunos, ovejas, cabras, búfalos, alces, antílopes, caribú y ciervos.

La expresión "péptidos señal," utilizada en el presente contexto se refiere a péptidos aislados de acuerdo con la presente invención obtenidos de cualquier especie, preferentemente microbiana, de cualquier origen, ya sea natural, sintético, semisintético o recombinante. Específicamente, se puede obtener un péptido señal de células metanógenas tales como células de Methanobrevibacter, especialmente células de M. ruminantium o M. smithii. Para la producción recombinante, se puede obtener un péptido señal de acuerdo con la presente invención de células microbianas o eucariotas por ejemplo, Escherichia, Streptomyces, Bacillus, Salmonella, levadura, células de insectos tales como Drosophila, células animales tales como células COS y CHO y células vegetales. Se ha de entender que cada referencia a un "péptido" en este contexto incluye la secuencia de longitud total (por ejemplo, SEC. ID. NO: 1-172), así como cualquier alteración, fragmento, derivado o variante de la misma.

Los términos "condiciones rigurosas" o "rigurosidad" utilizados en este contexto se refieren a las condiciones para hibridación de acuerdo con lo definido por el ácido nucleico, sal y temperatura. Estas condiciones son muy

- conocidas en la técnica y pueden ser modificadas para identificar o detectar secuencias de polinucleótidos idénticas o relacionadas. Remitirse, por ejemplo, a Sambrook, J. et al. (1989) *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Press, Plainview, NY, y Ausubel, F. M. et al. (1989) *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, NY. Numerosas condiciones equivalentes que comprenden baja o alta rigurosidad dependen de
- 5 factores tales como la longitud y la naturaleza de la secuencia (ADN, ARN, composición base), la naturaleza del objetivo (ADN, ARN, composición base), medio (en solución o inmovilizado en un sustrato sólido), la concentración de sales y otros componentes (por ejemplo, - formamida, sulfato de dextrano y/o polietilenglicol) y la temperatura de las reacciones (dentro de un rango de aproximadamente 5°C por debajo del punto de fusión de la sonda a
- 10 aproximadamente 20°C a 25°C por encima de la temperatura de fusión). Se puede variar uno o más factores para generar condiciones de baja o alta rigurosidad diferentes, aunque equivalentes a las condiciones antes enumeradas. El término "sujeto" incluye animales humanos y no humanos. Entre los animales no humanos se cuentan, aunque no a modo de limitación, aves y mamíferos tales como rumiantes y, en particular, ratones, conejos, gatos, perros, cerdos, ovejas, cabras, vacas y caballos.
- 15 Los términos "sustancialmente purificado" o "aislado" utilizados en este contexto se refieren a secuencias de ácido nucleico o aminoácidos que han sido separadas de su medio celular, recombinante o sintético y que están por lo menos 60 % libres, preferentemente 75 % libres y muy preferentemente por lo menos 90 % o por lo menos 99 % libres de otros componentes con los cuales están asociadas en un medio celular, recombinante o sintético.
- 20 En el presente contexto la "transformación," describe un proceso por el cual ingresa ADN exógeno y modifica una célula receptora. Esto puede tener lugar en condiciones naturales o artificiales empleando diversos métodos muy conocidos en la técnica. La transformación se puede basar en cualquier método conocido para la inserción de secuencias de ácidos nucleicos extrañas en una célula huésped procariota o eucariota. El método se elige basándose en el tipo de célula huésped que se está transformando y puede incluir, aunque no a modo de limitación,
- 25 la infección viral, electroporación, choque térmico, lipofección y bombardeo de partículas. Dichas células "transformadas" incluyen células transformadas de manera estable en las cuales el ADN insertado tiene la capacidad de replicarse ya sea como plásmido de replicación autónoma o como parte del cromosoma del huésped. También se incluyen las células que expresan transitoriamente el ADN o ARN insertado durante períodos de tiempo limitados. Una "variante" de un péptido o polipéptido, en el presente contexto, se refiere a una secuencia de aminoácidos que
- 30 está modificada por uno o más aminoácidos. Un polinucleótido variante está modificado por uno o más nucleótidos. Una variante puede dar lugar a cambios "conservadores", en los cuales un aminoácido sustituido tiene propiedades estructurales o químicas similares, por ejemplo el reemplazo de leucina por isoleucina. Más raramente, una variante puede dar lugar a cambios "no conservadores", por ejemplo el reemplazo de una glicina con un triptófano. Las
- 35 variaciones menores análogas también pueden incluir deleciones o inserciones de aminoácidos o ambas. Los lineamientos para determinar qué residuos aminoácidos se pueden sustituir, insertar o suprimir sin abolir la actividad biológica o inmunogénica se pueden encontrar empleando programas de computación muy conocidos en la técnica, por ejemplo el software LASERGENE (DNASTAR).
- 40 La invención abarca asimismo las variantes que retienen por lo menos una actividad biológica (por ejemplo, asociación a las células o permeabilización de las células) o actividad funcional del péptido o polipéptido. Una variante preferida es la que tiene por lo menos 80 %, y más preferentemente por lo menos 90 % de identidad de secuencia con una secuencia descrita. Una variante muy preferida es la que tiene por lo menos 95 %, por lo menos 97 %, por lo menos 98 %, por lo menos 99 %, por lo menos 99,5 %, por lo menos 99,8 % o por lo menos 99,9 % de
- 45 identidad de secuencia con una secuencia descrita en la presente. El porcentaje de identidad se determina alineando las dos secuencias a comparar de acuerdo con lo descrito más adelante, determinando el número de residuos idénticos incluidos en la porción alineada, dividiendo ese número por el número total de residuos de la secuencia de la invención (consultada) y multiplicando el resultado por 100. Un programa de alineación ventajoso es AlignX (Vector NTI).
- 50 El metano se produce en el intestino anterior de los rumiantes debido a los metanógenos que actúan como reductores terminales del carbono en el sistema del rumen. La vía de metanogénesis de múltiples etapas ha sido claramente dilucidada, principalmente por el estudio de metanógenos que no son del rumen, aunque no se han aclarado del todo las adaptaciones que permiten que los metanógenos se desarrollen y persistan en el rumen.
- 55 *Methanobrevibacter ruminantium* es un metanógeno primordial en los rumiantes de Nueva Zelanda. De acuerdo con lo aquí descrito, el proyecto de secuencia genómica *M. ruminantium* exhibe un tamaño de aproximadamente 3,0 Mb y un contenido de GC de 33,68%. Como hallazgo importante, se encontró que el genoma de *M. ruminantium* incluía secuencias peptídicas señal para usar en la selección como blanco y permeabilización de células. Por lo tanto, la invención abarca péptidos señal, incluyendo los que comprenden SEC. ID. NO: 1-172, así como polipéptidos que
- 60 comprenden estos péptidos y alteraciones, fragmentos, variantes y derivados de los mismos.

La invención abarca el uso de estos péptidos o polipéptidos para seleccionar como objetivo y permeabilizar células microbianas, especialmente células metanógenas. La invención abarca asimismo el uso de los péptidos o

polipéptidos para la inhibición del desarrollo o replicación de esas células. Los péptidos y polipéptidos de la presente invención se pueden expresar u utilizar en diversos ensayos para determinar su actividad biológica. Los péptidos y polipéptidos se pueden utilizar para la síntesis protocolos de aislamiento en gran escala y, por ejemplo para la producción comercial. Dichos péptidos y polipéptidos pueden ser utilizados para generar anticuerpos, aislar las correspondientes secuencias de aminoácidos y determinar cuantitativamente los niveles de las secuencias de aminoácidos.

Los polipéptidos de la presente invención pueden ser utilizados asimismo como composiciones, por ejemplo composiciones farmacéuticas, y como suplementos alimentarios, por ejemplo componentes del pienso para rumiantes. Los péptidos y polipéptidos de la presente invención pueden también ofrecer ventajas sanitarias. Por ejemplo, en los aspectos relacionados con la salud, se pueden utilizar inhibidores de metanógenos para restablecer la energía al sujeto, que normalmente pierde en forma de metano. En aspectos específicos, se pueden utilizar dispositivos para rumiantes en combinación con los péptidos, polipéptidos y composiciones (por ejemplo composiciones farmacéuticas y suplementos alimentarios) de la presente invención.

Los péptidos y polipéptidos de la presente invención comprenden por lo menos una secuencia seleccionada del grupo que consiste en: (a) péptidos o polipéptidos que comprenden por lo menos una secuencia de aminoácidos seleccionada del grupo que consiste en SEC. ID. NO: 1-172 o alteraciones, fragmentos, variantes o derivadas de las mismas, (b) péptidos o polipéptidos que comprenden un dominio funcional (por ejemplo, una región núcleo conservada descrita en la presente) de por lo menos una secuencia de aminoácidos seleccionada del grupo que consiste en SEC. ID. NO: 1-172 o alteraciones, fragmentos, variantes o derivadas de las mismas y (c) péptidos o polipéptidos que comprenden por lo menos un número especificado de residuos contiguos de por lo menos una secuencia de aminoácidos seleccionada del grupo que consiste en SEC. ID. NO: 1-172 alteraciones, fragmentos, variantes o derivadas de las mismas. En la presente se hace referencia a todas estas secuencias, en forma colectiva, como péptidos y polipéptidos de la presente invención. En una modalidad, la invención abarca un péptido o polipéptido aislado que comprende la secuencia de aminoácidos de por lo menos una de SEC. ID. NO: 1-172. La invención abarca además polinucleótidos que codifican por lo menos un péptido señal, que incluyen los correspondientes a SEC. ID. NO: 1-172, como así también los polipéptidos que comprenden estos péptidos y alteraciones, fragmentos, variantes o derivados de los mismos.

La invención abarca además el uso de estos polinucleótidos para la preparación de vectores de expresión y células huésped para seleccionar como objetivos y permeabilizar células microbianas, especialmente células metanógenas. La invención abarca asimismo el uso de los polinucleótidos para la inhibición del desarrollo o replicación de dichas células. Los polinucleótidos aislados de acuerdo con la presente invención también son de utilidad en el mapeo de genomas, en el mapeo físico y en la clonación de genes de bacterias más o menos relacionadas. Las sondas diseñadas mediante el uso de los polinucleótidos de la presente invención se pueden utilizar para detectar la presencia y examinar los patrones de expresión de genes en cualquier organismo que tenga secuencias de ADN y ARN suficientemente homólogas y sus células, empleando técnicas muy conocidas en el medio, tales como técnicas de transferencia por ranuras o análisis de micromatrices. Se pueden utilizar los cebadores diseñados utilizando los polinucleótidos de la presente invención para secuenciación y amplificaciones por PCR.

También se pueden utilizar los polinucleótidos de la presente invención en forma de composiciones, por ejemplo composiciones farmacéuticas, como suplementos alimentarios, por ejemplo componentes del pienso para rumiantes. Los polinucleótidos de la presente invención también ofrecen beneficios para la salud. En el caso de esos beneficios, los polinucleótidos se pueden presentar en forma de vectores de expresión o células huésped que comprenden vectores de expresión. En aspectos específicos, se pueden utilizar dispositivos para el rumen de liberación lenta en combinación con los polinucleótidos, vectores, células huésped y composiciones (por ejemplo composiciones farmacéuticas y suplementos alimentarios) de la presente invención.

Los polinucleótidos de la presente invención comprenden por lo menos una secuencia seleccionada del grupo que consiste en: (a) secuencias que comprenden por lo menos una secuencia de aminoácidos seleccionada del grupo que consiste en SEC. ID. NO: 1-172 o alteraciones, fragmentos, variantes o derivados de las mismas; (b) complementos, secuencias inversas y complementos inversos de una secuencia codificadora correspondiente a por lo menos una secuencia de aminoácidos seleccionada del grupo que consiste en SEC. ID. NO: 1-172 o alteraciones, fragmentos, variantes o derivados de las mismas; (c) marcos de lectura abiertos contenidos en la secuencia codificadora correspondiente a por lo menos una secuencia de aminoácidos seleccionada del grupo que consiste en SEC. ID. NO: 1-172 o sus alteraciones, fragmentos, variantes o derivados; (d) dominios funcionales (por ejemplo, regiones núcleo conservadas descritas en la presente) de una secuencia codificadora correspondiente a por lo menos una secuencia de aminoácidos seleccionada del grupo que consiste en SEC. ID. NO: 1-172 o alteraciones, fragmentos, variantes o derivados de las mismas y (e) secuencias que comprenden por lo menos un número especificado de residuos contiguos de una secuencia codificadora correspondiente a por lo menos una secuencia de aminoácidos seleccionada del grupo que consiste en SEC. ID. NO: 1-172 o alteraciones, fragmentos, variantes o derivados de las mismas. También se dan a conocer sondas y cebadores oligonucleotídicos. En el

contexto del presente documento, se hace referencia a todos estos polinucleótidos y sondas y cebadores oligonucleotídicos, en forma colectiva, como polinucleótidos de la presente invención. En una modalidad, la invención abarca un polinucleotido aislado que comprende una secuencia codificadora para por lo menos una secuencia de aminoácidos seleccionada del grupo que consiste en SEC. ID. NO: 1–172.

5

Los expertos en la técnica sabrán apreciar que, como resultado de la degeneración del código genético, se puede producir una multitud de secuencias nucleotídicas que codifican los péptidos de la presente invención, algunos con mínima homología con las secuencias nucleotídicas de cualquier gen conocido y de origen natural. Por consiguiente, la invención contempla todas y cada una de las variaciones posibles de las secuencias nucleotídicas que se podrían generar seleccionando combinaciones basándose en elecciones de los codones posibles. Estas combinaciones se generan de acuerdo con los tripletes del código genético standard aplicado a las secuencias de origen natural y todas dichas variaciones se deben considerar específicamente descritas.

10

Las secuencias nucleotídicas que codifican péptidos o polipéptidos señalan o sus secuencias modificadas son capaces preferentemente de hibridarse a la secuencia nucleotídica de la secuencia de origen natural en condiciones de rigurosidad seleccionadas de manera Sin embargo, puede ser ventajoso producir secuencias de nucleótidos que codifiquen un péptido o sus derivados, que posea un uso de codones sustancialmente diferente. Los codones que se pueden seleccionar para aumentar la velocidad a la que tiene lugar la expresión del péptido en un huésped procarionta o eucariota específico de acuerdo con la frecuencia con la cual los codones específicos son utilizados por el huésped. Por ejemplo, se pueden optimizar los codones para la expresión en *E. coli*, por ejemplo, la provista por SEC. ID. NO: 342–533. Otras razones para alterar sustancialmente la secuencia nucleotídica que codifica péptidos y sus derivados sin alterar las secuencias de aminoácidos incluyen la producción de transcripciones de ARN con más propiedades convenientes, como por ejemplo una vida media más prolongada que las transcripciones producidas a partir de la secuencia de origen natural.

20

La invención abarca asimismo la producción de secuencias de ADN o fragmentos de las mismas, que codifican los péptidos o polipéptidos o sus secuencias modificadas, completamente por química sintética. Después de la producción, la secuencia sintética se puede insertar en cualquiera de los numerosos vectores de expresión y sistemas celulares disponibles empleando reactivos que son muy cot reactivos que son muy conocidos en la técnica.

30

Más aun, se puede utilizar la química de síntesis para introducir mutaciones en una secuencia que codifica un péptido o polipéptido o cualquier alteración, variante, derivado o fragmento de la misma. También están cubiertas por la invención las secuencias de polinucleótidos con capacidad de hibridarse a las secuencias nucleotídicas reivindicadas y, en particular, las expuestas en SEC. ID. NO: 173–341 o 342–533 o sus complementos, en diversas condiciones de rigurosidad de acuerdo con lo descrito por Wahl, G. M. y S. L. Berger (1987; *Methods Enzymol.* 152:399–407) y Kimmel, A. R. (1987; *Methods Enzymol.* 152:507–511).

35

Los métodos para la secuenciación de ADN son muy conocidos y por lo general están disponibles en la técnica y se los puede utilizar para poner en práctica cualquiera de las modalidades de la presente invención. Los métodos pueden emplear enzimas tales como el fragmento Klenow de la ADN polimerasa I, SEQUENASE (U.S. Biochemical Corp, Cleveland, OH), polimerasa Taq (Perkin Elmer), polimerasa T7 termoestable de Amersham Pharmacia Biotech (Piscataway, NJ) o combinaciones de polimerasas y exonucleasas para corrección de pruebas tales como las que se encuentran en el Sistema de Amplificación ELONGASE comercializado por Life Technologies (Gaithersburg, MD). De preferencia, el proceso se automatiza con máquinas tales como el Hamilton Micro Lab 2200 (Hamilton, Reno, NV), el Cicladador Térmico Peltier (PTC200; MJ Research, Watertown, MA) los secuenciadores de ADN ABI Catalyst y 373 y 377 (Perkin Elmer) o el Genome Sequencer 20TM (Roche Diagnostics).

45

Las secuencias de ácidos nucleicos que codifican los péptidos se pueden prolongar utilizando una secuencia nucleotídica parcial y empleando diversos métodos conocidos en la técnica para la detección de secuencias de la parte anterior tales como promotores y elementos regulatorios. Por ejemplo, un método que se puede emplear, PCR de “sitios de restricción”, utiliza cebadores universales para recuperar una secuencia desconocida adyacente a un locus conocido (Sarkar, G. (1993) *PCR Methods Applic.* 2:318–322). En particular, en primer lugar se amplifica ADN genómico en presencia de un cebador para una secuencia ligante y un cebador específico para la región conocida. A continuación se somete a las secuencias amplificadas a una segunda ronda de PCR con el mismo cebador ligante y otro cebador específico interno con respecto al primero. Los productos de cada ronda de PCR son transcritos con una ARN polimerasa adecuada y secuenciados utilizando transcriptasa inversa.

55

Se pueden utilizar los sistemas de electroforesis capilar que existen en el comercio para analizar el tamaño o confirmar la secuencia nucleotídica de los productos de la secuenciación o la PCR. Específicamente, la secuenciación capilar puede emplear polímeros fluyentes para la separación por electroforesis, cuatro tinturas fluorescentes diferentes (una por cada nucleótido) que son activadas por láser y la detección de las longitudes de onda emitidas por una cámara del dispositivo acoplada con carga. Se puede convertir la salida/intensidad de la luz a una señal eléctrica utilizando el software apropiado (por ejemplo, GENOTYPER y Sequence NAVIGATOR, Perkin Elmer) y se puede controlar por computadora la totalidad del proceso, desde la carga de las muestras hasta el

60

análisis informático y la visualización de los datos electrónicos. La electroforesis capilar es especialmente preferible para la secuenciación de pequeños trozos de ADN que pudieran estar presentes en cantidades limitadas en una muestra específica.

- 5 En otra modalidad de la presente invención, se pueden utilizar polinucleótidos o fragmentos de los mismos que codifican péptidos o polipéptidos en moléculas de ADN recombinante par dirigir la expresión de los péptidos, polipéptidos o secuencias modificadas de los mismos, en células huésped apropiadas. Debido a la degeneración inherente del código genético, se pueden producir otras secuencias de ADN que codifican secuencias de aminoácidos sustancialmente iguales o funcionalmente equivalentes, y se pueden utilizar estas secuencias para clonar y expresar péptidos o polipéptidos señal. Las secuencias de nucleótidos de la presente invención se pueden modificar utilizando métodos generalmente conocidos en la técnica a fin de alterar las secuencias de aminoácidos por una variedad de razones, entre las que se incluyen, aunque no a modo de limitación, alteraciones que modifican la clonación, procesamiento y/o expresión del producto génico. Se puede recurrir al reordenamiento de ADN por fragmentación al azar y el reensamble por PCR de los fragmentos génicos y oligonucleótidos sintéticos para modificar las secuencias nucleotídicas. Por ejemplo, se puede emplear la mutagénesis dirigida a un sitio específico para insertar nuevos sitios de restricción, modificar patrones de glicosilación, cambiar la preferencia por codones, introducir mutaciones y así sucesivamente.

En otra modalidad de la presente invención, se pueden ligar secuencias de ácidos nucleicos naturales, modificadas o recombinantes que codifican los péptidos o polipéptidos a una secuencia heteróloga para codificar una proteína de fusión. Por ejemplo, puede ser de utilidad codificar una secuencia quimérica que pueda ser reconocida por un anticuerpo existente en el comercio. También se puede diseñar una proteína de fusión que contenga un sitio de escisión situado entre el péptido o polipéptido de la presente invención y la secuencia proteica heteróloga, de manera que el péptido o polipéptido se pueda escindir y pueda ser purificada aparte de la porción heteróloga.

En otra modalidad, se pueden sintetizar secuencias que codifiquen péptidos o polipéptidos, en todo o en parte, utilizando métodos químicos muy conocidos en la técnica (Ver Caruthers, M. H. et al. (1980) Nucl. Acids Res. Symp. Ser. 215–223, Horn, T. et al. (1980) Nucl. Acids Res. Symp. Ser. 225–232). Por otro lado, se puede producir el péptido o polipéptido en sí utilizando métodos químicos para sintetizar la secuencia de aminoácidos o un fragmento de la misma. Por ejemplo, se puede realizar la síntesis peptídica utilizando diversas técnicas en fase sólida (Roberge, J. Y. et al. (1995) Science 269:202–204; Merrifield J. (1963) J. Am. Chem. Soc. 85:2149–2154) y se puede obtener la síntesis automática, por ejemplo, utilizando el Sintetizador de Péptidos ABI 431A (Perkin Elmer). Se pueden sintetizar químicamente varios fragmentos de péptidos o polipéptidos por separado y combinarlos empleando métodos químicos para producir la molécula completa.

El péptido o polipéptido recién sintetizado puede ser aislado por cromatografía líquida de alto rendimiento (por ejemplo, Creighton, T. (1983) Proteins Structures and Molecular Principles, WH Freeman and Co., New York, NY). La composición de los péptidos o polipéptidos sintéticos se puede confirmar mediante análisis de aminoácidos o secuenciación (por ejemplo, el procedimiento de degradación de Edman; Creighton, supra). Además, la secuencia de aminoácidos del péptido o polipéptido o cualquier parte de la misma se puede modificar durante la síntesis directa y/o combinar utilizando métodos químicos con secuencias de otras proteínas o cualquier parte de las mismas, para producir una molécula modificada.

Para expresar un péptido biológicamente activo, las secuencias nucleotídicas que codifican el péptido o equivalentes funcionales, se pueden insertar las secuencias que codifican el péptido o equivalentes funcionales en un vector de expresión apropiado, es decir, un vector que contiene los elementos necesarios para la transcripción y traducción de la secuencia codificadora insertada. Se pueden utilizar métodos que son muy conocidos por Los expertos en la técnica para construir vectores de expresión que contienen secuencias que codifican el péptido y los elementos de transcripción y traducción apropiados. Estos métodos incluyen técnicas de ADN recombinante in vitro, técnicas sintéticas y recombinación genética in vivo. Dichas técnicas han sido descritas por Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview, NY, y Ausubel, F. M. et al. (1989) Current Protocols in Molecular Biology, John Wiley & Sons, New York, NY.

Se puede utilizar una variedad de sistemas vector de expresión /huésped para que contengan y expresen secuencias que codifican los péptidos de la presente invención. Entre estos se cuentan, aunque no a modo de limitación, microorganismos tales como bacterias transformadas con vectores de expresión de bacteriófagos recombinantes, plásmidos o cósmido ADN, levadura transformada con vectores de expresión en levadura, sistemas celulares de insectos infectados con vectores de expresión en virus (por ejemplo, baculovirus), sistemas de células vegetales transformados con vectores de expresión virales (por ejemplo, el virus del mosaico de la coliflor, CaMV, el virus del mosaico del tabaco, TMV) o con vectores de expresión bacteriana (por ejemplo, los plásmidos Ti o pBR322); o sistemas celulares animales. En el caso de las bacterias, los plásmidos provechosos incluyen los plásmidos pET, pRSET, pTrcHis2, y pBAD de Invitrogen, pET y pCDF de Novagen, y los plásmidos DirectorTM de Sigma–Aldrich. En el caso de los metanógenos, entre los plásmidos útiles se cuentan, aunque no a modo de

limitación, pME2001, pMV15 y pMP1. Específicamente, se puede utilizar *Escherichia coli* con el vector de expresión pET. La invención no está limitada por el vector de expresión ni la célula huésped empleada.

Los “elementos de control” o “secuencias regulatorias” son las regiones no traducidas del vector –potenciadores, promotores, regiones sin traducir 5’ y 3’ – que interactúan con las proteínas celulares del huésped para llevar a cabo la transcripción y la traducción. Dichos elementos pueden variar en cuanto a su potencia y especificidad. Dependiendo del sistema vector y el huésped utilizado, se puede utilizar cualquier número de elementos de transcripción y traducción adecuados, incluyendo promotores constitutivos e inducibles. Por ejemplo, cuando se realiza la clonación en sistemas bacterianos, se pueden utilizar promotores inducibles tales como el promotor híbrido lacZ del fagémido BLUESCRIPT (Stratagene, LaJolla, CA) o el plásmido pSPORT1 (Life Technologies) y demás. En células de insectos se puede utilizar el promotor de polihedrina de baculovirus. Los promotores o potenciadores derivados de los genomas de células vegetales (por ejemplo, genes termosensibles, RUBISCO y de almacenamiento de proteínas) o de virus de las plantas (por ejemplo, promotores virales o secuencias quí) se pueden clonar en el vector.

En los sistemas bacterianos, se puede seleccionar un número de vectores de expresión dependiendo del uso pretendido del péptido. Por ejemplo, cuando se necesitan grandes cantidades de péptido, se pueden utilizar vectores que dirigen un alto nivel de expresión de proteínas de fusión que son fácilmente purificadas. Entre dichos vectores se incluyen, aunque no a modo de limitación, los vectores de clonación y expresión multifuncionales en *E. coli* tales como BLUESCRIPT (Stratagene), en los cuales se puede ligar la secuencia que codifica un péptido a un vector en marco con las secuencias para la Met amino-terminal y los 7 residuos subsiguientes de la β -galactosidasa de manera que se produzca una proteína híbrida, vectores pIN (Van Heeke, G. y S. M. Schuster (1989) *J. Biol. Chem.* 264:5503–5509); y demás.

También se pueden utilizar vectores pGEX (Promega, Madison, WI) para expresar péptidos extraños como proteínas de fusión con glutatión S-transferasa (GST). En general, dichas proteínas de fusión son solubles y pueden ser fácilmente purificadas de células lisadas mediante absorción a perlas de glutatión-agarosa para luego proceder a la elución en presencia de glutatión libre. Las proteínas preparadas en dichos sistemas pueden ser diseñadas de manera que incluyan sitios de escisión de heparina, trombina o factor Xa proteasa de manera que el péptido de interés clonado se pueda liberar de la porción GST a voluntad. En la levadura *Saccharomyces cerevisiae*, se puede utilizar un número de vectores que contienen promotores constitutivos o inducibles tales como factor alfa, alcohol oxidasa y PGH. Para encontrar reseñas, remitirse a Ausubel et al. (supra) and Grant et al. (1987) *Methods Enzymol.* 153:516–544.

También se pueden utilizar señales de iniciación específicas para obtener una traducción más eficiente de las secuencias que codifican los péptidos de la presente invención. Dichas señales incluyen el codón de iniciación ATG y las secuencias adyacentes. En los casos en que se insertan secuencias que codifican un péptido, su codón de iniciación y secuencias de la porción anterior en el vector de expresión apropiado, pueden no ser necesarias otras señales de control de la transcripción o la traducción. Sin embargo, en los casos en que sólo se inserta una secuencia codificadora o un fragmento de la misma, se deben incluir señales exógenas de control de la traducción que incluyan el codón de iniciación ATG. Más aun, el codón de iniciación debe estar en el marco de lectura correcto para garantizar la traducción de la totalidad del inserto. Los elementos de traducción y codones de iniciación exógenos pueden ser de diversos orígenes, tanto naturales como sintéticos. Se puede potenciar la eficiencia de la expresión mediante la inclusión de potenciadores que son apropiados para el sistema celular específico que se utiliza, como por ejemplo los descritos en la literatura (Scharf, D. et al. (1994) *Results Probl. Cell Differ.* 20:125–162). Además, se puede elegir una cepa de célula huésped por su capacidad para modular la expresión de las secuencias insertadas o procesar el péptido o polipéptido expresado en la forma pretendida. Dichas modificaciones de la secuencia incluyen, aunque no a modo de limitación, la acetilación, carboxilación, glicosilación, fosforilación, lipidación y acilación. También se puede hacer uso del procesamiento postraducción que escinde una forma “prepro” del péptido o polipéptido para facilitar la correcta inserción, plegamiento y/o función. Las diferentes células huésped que tienen una maquinaria celular específica y mecanismo característicos para actividades post-traducción se pueden obtener en la American Type Culture Collection (ATCC; Bethesda, MD) y pueden ser elegidas para garantizar la modificación y el procesamiento correctos de la secuencia. Entre las células huésped específicas se cuentan, aunque no a modo de limitación, células de metanógenos tales como células de *Methanobrevibacter*, especialmente células de *M. ruminantium* o *M. smithii*. Las células huésped de interés incluyen, por ejemplo, *Rhodotorula*, *Aureobasidium*, *Saccharomyces*, *Sporobolomyces*, *Pseudomonas*, *Erwinia* y *Flavobacterium*; u otros organismos tales como *Escherichia*, *Lactobacillus*, *Bacillus*, *Streptomyces*, y demás. Entre las células huésped específicas se cuentan *Escherichia coli*, que es especialmente adecuada para usar en la presente invención, *Saccharomyces cerevisiae*, *Bacillus thuringiensis*, *Bacillus subtilis*, *Streptomyces lividans*, y demás.

Hay varias técnicas para introducir ácidos nucleicos en células eucariotas cultivadas in vitro. Entre éstas se incluyen los métodos químicos (Felgner et al., *Proc. Natl. Acad. Sci., USA*, 84:7413–7417 (1987); Bothwell et al., *Methods for Cloning and Analysis of Eukaryotic Genes*, Eds., Jones and Bartlett Publishers Inc., Boston, Mass. (1990), Ausubel et

al., Short Protocols in Molecular Biology, John Wiley and Sons, New York, NY (1992) y Farhood, *Annal. NY Acad. Sci.*, 716:23 34 (1994)), el uso de protoplastos (Bothwell, *supra*) o pulsos eléctricos (Vatteroni et al., *Mutn. Res.*, 291:163 169 (1993); Sabelnikov, *Prog. Biophys. Mol. Biol.*, 62: 119 152 (1994); Bothwell et al., *supra*; y Ausubel et al., *supra*), el uso de virus atenuados (Davis et al., *J. Virol.* 1996, 70(6), 3781 3787; Brinster et al. *J. Gen. Virol.* 2002, 5 83(Pt 2), 369 381; Moss, *Dev. Biol. Stan.*, 82:55 63 (1994); and Bothwell et al., *supra*), así como métodos físicos (Fynan et al., *supra*; Johnston et al., *Meth. Cell Biol.*, 43(Pt A):353 365 (1994); Bothwell et al., *supra* y Ausubel et al., *supra*).

La aplicación favorable de ácidos nucleicos a tejidos animales se puede obtener por medio de liposomas catiónicos
 10 (Watanabe et al., *Mol. Reprod. Dev.*, 38:268 274 (1994)), inyección directa de ADN o ARN desnudo en el tejido muscular del animal (Robinson et al., *Vacc.*, 11:957 960 (1993); Hoffman et al., *Vacc.* 12:1529 1533; (1994); Xiang et al., *Virol.*, 199:132 140 (1994); Webster et al., *Vacc.*, 12:1495 1498 (1994); Davis et al., *Vacc.*, 12:1503 1509 (1994); Davis et al., *Hum. Molec. Gen.*, 2:1847 1851 (1993); Dalemans et al. *Ann NY Acad. Sci.* 1995, 772, 255 256. Conry, et al. *Cancer Res.* 1995, 55(7), 1397–1400), y embriones (Naito et al., *Mol. Reprod. Dev.*, 39:153 161 (1994); y
 15 Burdon et al., *Mol. Reprod. Dev.*, 33:436 442 (1992)), inyección intramuscular de vacunas de autorreplicación de ARN (Davis et al., *J Virol* 1996, 70(6), 3781 3787; Balasuriya et al. *Vaccine* 2002, 20(11 12), 1609 1617) o inyección intradérmica de ADN utilizando la tecnología de “pistola de genes” (Johnston et al., *supra*).

En la técnica se conoce una variedad de protocolos para la detección y medición de la expresión de los péptidos o
 20 polipéptidos de la presente invención, ya sea utilizando anticuerpos policlonales o monoclonales específicos para la proteína. Entre los ejemplos se cuentan el ensayo de inmunoabsorbencia ligado a enzimas (ELISA), radioinmunoensayo (RIA), y clasificación celular activada por fluorescencia (FACS). Se puede utilizar un inmunoensayo de dos sitios, basado en anticuerpos monoclonales con anticuerpos monoclonales reactivos a dos
 25 epítopes que no interfieran del péptido o polipéptido, aunque también se puede utilizar un ensayo de unión competitiva. Estos y otros ensayos han sido descritos, entre otros, por Hampton, R. et al. (1990; *Serological Methods, a laboratory Manual*, APS Press, St Paul, MN) and Maddox, D. E. et al. (1983; *J. Exp. Med.* 158:1211–1216).

Los expertos en la técnica conocen una amplia variedad de marcadores y técnicas de conjugación que pueden
 30 utilizar en diversos ensayos de ácidos nucleicos y aminoácidos. Los medios para la producción de sondas marcadas para hibridación o PCR para la detección de secuencias relacionadas con los polinucleótidos incluyen oligomarcación, traducción de mellas, marcación de extremos o amplificación de PCR utilizando un nucleótido marcado. Por otro lado, se pueden clonar las secuencias que codifican el péptido o los polipéptidos que puedan comprender este péptido o cualquier secuencia modificada del mismo en un vector para la producción de una sonda
 35 de ARNm. Dichos vectores son conocidos en la técnica, existen en el comercio y se los puede utilizar para sintetizar sondas de ARN in vitro mediante la adición de una ARN polimerasa apropiada tal como T7, T3 o SP6 y nucleótidos marcados. Estos procedimientos se pueden llevar a cabo utilizando una variedad de kits existentes en el comercio, de Amersham Pharmacia Biotech, Promega t US Biochemical. Las moléculas marcadoras o marcas adecuadas que se pueden utilizar para facilitar la detección incluyen radionúclidos, enzimas, agentes fluorescentes,
 40 quimioluminiscentes o cromogénicos, como así también sustratos, cofactores, inhibidores, partículas magnéticas y demás.

Se pueden cultivar vectores de expresión o células huésped transformadas con vectores de expresión en condiciones adecuadas para la expresión y recuperación del péptido o polipéptido del cultivo. El cultivo puede
 45 comprender componentes para la expresión in vitro o in vivo. Los componentes para la expresión in vitro incluyen los correspondientes a los lisados de reticulocitos de conejo, lisados de *E. coli* y extractos de germen de trigo, por ejemplo, los sistemas Expressway™ o RiPs de Invitrogen, los sistemas Genelator™ de iNtRON Biotechnology, los sistemas EcoPro™ o STP3™ de Novagen, los sistemas TNT® Quick Coupled de Promega y los sistemas EasyXpress de QIAGEN. El péptido o polipéptido producido por el cultivo puede ser secretado o contenido en el interior de la célula dependiendo de la secuencia y/o el vector empleado. Como comprenderán Los expertos en la
 50 técnica, los vectores de expresión que codifican los péptidos o polipéptidos son diseñados preferentemente de manera que contengan secuencias señal que dirigen la secreción del péptido a través de la membrana de una célula procariota o eucariota.

Otras construcciones pueden incluir un dominio aminoácido que facilite la purificación del péptido o polipéptido. Dichos dominios incluyen, aunque no a modo de limitación, péptidos quelantes de metal tales como moléculas de
 55 histidina–triptófano (por ejemplo, 6X–HIS (SEQ ID NO: 514)) que permiten la purificación en metales inmovilizados, dominios de proteína A que permiten la purificación en inmunoglobulina inmovilizada y el dominio utilizado en el sistema de purificación por extensión/afinidad FLAG® (Immunex Corp., Seattle, WA). Las marcas de epítopes útiles
 60 incluyen 3XFLAG®, HA, VSV–G, V5, HSV, GST, GFP, MBP, GAL4, y β–galactosidasa. Entre los plásmidos ventajosos se cuentan los que comprenden una marca de biotina (por ejemplo los plásmidos PinPoint™ de Promega), proteína de unión a calmodulina (por ejemplo, plásmidos pCAL de Stratagene), péptido de unión a estreptavidina (por ejemplo, los plásmidos InterPlay™ de Stratagene), una marca c–myc o FLAG® (por ejemplo,

plásmidos de inmunoprecipitación de Sigma–Aldrich) o una marca de histidina (por ejemplo, los plásmidos QIAExpress de QIAGEN).

Para facilitar la purificación, se puede utilizar una secuencia ligante escindible, como las específicas para el Factor Xa o enteroquinasa (Invitrogen, San Diego, CA). Por ejemplo, el vector puede incluir uno o más ligantes entre el dominio de purificación y el péptido o polipéptido. En un aspecto, el vector de expresión puede dar lugar a la expresión de una proteína de fusión que comprende el péptido o polipéptido de la presente invención y un ácido nucleico que codifica 6 residuos de histidina (SEQ ID NO: 514) que precede a un sitio de escisión de tiorredoxina o enteroquinasa. Los residuos de histidina facilitan la purificación en IMAC (cromatografía de afinidad de iones metal inmovilizados según lo descrito por Porath, J. et al. (1992) *Prot. Exp. Purif.* 3: 263–281) en tanto que el sitio de escisión de enteroquinasa ofrece un medio para la purificación del péptido o polipéptido de la proteína de fusión. Se puede encontrar una explicación de los vectores con contenido de proteínas de fusión en el trabajo de Kroll, D. J. et al. (1993; *DNA Cell Biol.* 12:441–453).

Los anticuerpos de la presente invención se pueden producir empleando métodos generalmente conocidos en la técnica, por ejemplo, para usar en técnicas de purificación o diagnóstico. En particular, se pueden utilizar péptidos, polipéptidos o polinucleótidos purificados para producir anticuerpos de acuerdo con protocolos generalmente conocidos. Dichos anticuerpos pueden incluir, aunque no a modo de limitación, anticuerpos policlonales, monoclonales, quiméricos y monocatenarios, fragmentos Fab y fragmentos producidos por una biblioteca de expresión de Fab. Son especialmente preferibles los anticuerpos neutralizantes (es decir, los que inhiben la función) para usar con la presente invención.

Para la producción de anticuerpos, se pueden inmunizar diversos huéspedes, incluyendo cabras, conejos, ratas, ratones, humanos y otros, mediante inyección con un péptido, polipéptido, polinucleótido o cualquier fragmento de los mismos que tenga propiedades inmunogénicas. Dependiendo de la especie del huésped, se pueden utilizar diversos adyuvantes para aumentar la respuesta inmunológica. Dichos adyuvantes incluyen, aunque no a modo de limitación, el de Freund, geles minerales tales como hidróxido de aluminio y sustancias tensioactivas tales como lisolecitina, polioles plurónicos, polianiones, péptidos, emulsiones oleosas, hemocianina de lapa californiana y dinitrofenol. Entre los adyuvantes empleados en humanos, son especialmente preferibles BCG (bacilos de Calmette–Guerin) y *Corynebacterium parvum*.

Es preferible que los péptidos, polipéptidos o fragmentos utilizados para inducir anticuerpos tengan una secuencia de aminoácidos que comprende por lo menos cinco aminoácidos y más preferentemente por lo menos 10 aminoácidos. También es preferible que sean idénticos a una porción de la secuencia de aminoácidos de la proteína natural y pueden contener la secuencia de aminoácidos completa de una molécula pequeña de origen natural. Se pueden fusionar tramos cortos de aminoácidos a los de otra proteína tal como hemocianina de lapa californiana y producir anticuerpos contra la molécula quimérica.

Los anticuerpos monoclonales se pueden preparar empleando cualquier técnica que dé lugar a la producción de moléculas de anticuerpos por líneas celulares continuas en cultivo. Estas incluyen, aunque no a modo de limitación, la técnica de hibridomas, la técnica de hibridoma de células B humanas y la técnica de EBV–hibridoma (Kohler, G. et al. (1975) *Nature* 256:495–497; Kozbor, D. et al. (1985) *J. Immunol. Methods* 81:31–42; Cote, R. J. et al. (1983) *Proc. Natl. Acad. Sci.* 80:2026–2030; Cole, S. P. et al. (1984) *Mol. Cell Biol.* 62:109–120). También se pueden producir anticuerpos induciendo la producción in vivo en la población de linfocitos o seleccionando bibliotecas de inmunoglobulina o paneles de reactivos de unión sumamente específicos de acuerdo con lo descrito en la literatura (Orlandi, R. et al. (1989) *Proc. Natl. Acad. Sci.* 86:3833–3837; Winter, G. et al. (1991) *Nature* 349:293–299).

Además, se pueden emplear técnicas para la producción de “anticuerpos quiméricos”, por ejemplo la combinación de genes de anticuerpos para obtener una molécula con la especificidad por el antígeno y actividad biológica apropiadas (Morrison, S. L. et al. (1984) *Proc. Natl. Acad. Sci.* 81:6851–6855; Neuberger, M.S. et al. (1984) *Nature* 312:604–608; Takeda, S. et al. (1985) *Nature* 314:452–454). Por otro lado, las técnicas descritas para la producción de anticuerpos de cadena simple se pueden adaptar, empleando métodos conocidos en la técnica, a la producción de anticuerpos específicos de cadena simple. Se pueden generar anticuerpos con especificidad relacionada, aunque de composición idiotípica diferenciada, mediante reordenamiento de cadena a partir de bibliotecas de inmunoglobulinas por combinación (Burton D. R. (1991) *Proc. Natl. Acad. Sci.* 88:11120–3).

Los expertos en la técnica a la cual la presente invención pertenece sabrán comprender los términos “diacuerpos” y “triacuerpos”. Estas son moléculas que comprenden un dominio de cadena pesada variable (VH) conectado a un dominio de cadena liviana variable (VL) por un corto ligante peptídico que es demasiado corto para permitir el apareamiento entre los dos dominios en la misma cadena. Esto promueve el apareamiento con los dominios complementarios de una o más cadenas adicionales y fomenta la formación de moléculas dimericas o triméricas con dos o más sitios funcionales de unión al antígeno. Las moléculas de anticuerpos así obtenidas pueden ser mono-específicas o multi-específicas (por ejemplo biespecíficas en el caso de los diacuerpos). Dichas moléculas de anticuerpos pueden ser generadas a partir de dos o más anticuerpos utilizando metodología standard en la técnica a

la cual la invención pertenece, por ejemplo de acuerdo con lo descrito por Todorovska et al. (Design and application of diabodies, triabodies and tetrabodies for cancer targeting. *J. Immunol. Methods*. 2001 Feb 1;248(1-2):47-66).

También se pueden generar fragmentos de anticuerpos que contienen sitios de unión específicos. Por ejemplo, dichos fragmentos incluyen, aunque no a modo de limitación, fragmentos F(ab')₂ que se pueden producir mediante
5 digestión con pepsina de la molécula de anticuerpo y los fragmentos Fab que se pueden generar reduciendo los puentes de disulfuro de los fragmentos F(ab')₂. Por otro lado, se pueden construir bibliotecas de expresión de Fab para permitir la rápida y fácil identificación de los fragmentos Fab monoclonales con la especificidad pretendida (Huse, W. D. et al. (1989) *Science* 254:1275-1281).

- 10 Se pueden utilizar diversos inmunoensayos para el análisis para la identificación de anticuerpos con especificidad de unión. Numerosos protocolos para la unión competitiva o ensayos inmunoradiométricos utilizando anticuerpos policlonales o monoclonales con especificidades establecidas son muy conocidos en la técnica. Dichos inmunoensayos implican típicamente la medición de la formación de complejos entre un péptido, polipéptido o polinucleótido y su anticuerpo específico. Es preferible un inmunoensayo de dos sitios, basado en anticuerpos
15 monoclonales que utilizan anticuerpos monoclonales reactivos para dos epítopes no interferentes, aunque también se puede emplear un ensayo de unión competitiva (Maddox, supra).

Los péptidos señal descritos en la presente tienen la capacidad de penetrar en las células y, por lo tanto, son útiles como moléculas portadoras para la aplicación de moléculas inhibitorias a las células microbianas. La química para el
20 acoplamiento de los compuestos a aminoácidos está muy desarrollada y se podría ligar un número de tipos de moléculas diferentes al péptido señal. Los métodos de acoplamiento más comunes se basan en la presencia de grupos amino libres (alfa-amino o Lys), sulfhidrilo (Cys) o ácido carboxílico (Asp, Glu o alfa-carboxilo). Se pueden utilizar métodos de acoplamiento para ligar el péptido al inhibidor celular por medio del residuo carboxi- o amino-terminal. En algunos casos, una secuencia incluye múltiples residuos que pueden reaccionar con la química elegida.

- 25 Esto se puede utilizar para producir multímeros, que comprenden más de un inhibidor celular. Por otro lado, el péptido o polipéptido puede ser acertado o elegido de manera que los residuos reactivos estén localizados en el término amino o el término carboxilo de la secuencia.

Por ejemplo, se puede incorporar específicamente una molécula reportera tal como fluoresceína en un residuo de
30 lisina (Ono et al., 1997) utilizando N- α -Fmoc-N ϵ -1-(4.4-dimetil-2,6 dioxociclohex-1-ilideno-3-metilbutil)-L-lisina durante la síntesis del péptido. Después de la síntesis, se pueden acoplar ésteres de succinimidilo de 5- y 6-carboxifluoresceína, después de retirar el 4.4-dimetil-2,6 dioxociclohex-1-ilideno mediante tratamiento con hidrazina. Por lo tanto, el acoplamiento de una molécula inhibitoria el péptido o polipéptido señal se puede realizar mediante la inclusión de un residuo de lisina en la secuencia permeabilizante y luego la reacción con un inhibidor
35 celular adecuadamente derivado.

También se puede utilizar el método de acoplamiento de EDC (clorhidrato de 1-etil-3-(3-dimetilaminopropil) carbodiimida) o el de carbodiimida. Las carbodiimidias pueden activar los grupos carboxílico de cadena lateral del ácido aspártico y glutámico, así como el grupo carboxilo terminal para tornarlos sitios reactivos para el acoplamiento
40 con aminas primarias. Los péptidos activados se mezclan con el inhibidor celular para producir el conjugado final. Si se activa primero el inhibidor celular, el método EDC acopla el inhibidor celular a través de la alfa amina N-terminal y posiblemente a través de la amina de la cadena lateral de Lys, en caso de estar presente en la secuencia.

El éster de m-maleimidobenzoil-N-hidroxisuccinimida (MBS) es un reactivo heterobifuncional que se puede utilizar para ligar péptidos a inhibidores celulares por medio de cisteínas. El acoplamiento tiene lugar con el grupo tiol de los
45 residuos de cisteína. Si la secuencia elegida no contiene Cys, es común colocar un residuo Cys en el término N o C para obtener una ligadura sumamente controlada del péptido al inhibidor celular. Para fines de síntesis, puede ser conveniente colocar la cisteína en el término N del péptido. El MBS es particularmente adecuado para usar con la presente invención.

- 50 Se puede utilizar glutaraldehído como reactivo de acoplamiento bifuncional que liga a dos compuestos a través de sus grupos amino. El glutaraldehído otorga un espaciador sumamente flexible entre el péptido y el inhibidor celular para una presentación favorable. El glutaraldehído es un compuesto muy reactivo y reacciona con Cys, Tyr e His en un grado limitado. El método de acoplamiento con glutaraldehído es de particular utilidad cuando un péptido contiene sólo un grupo amino libre único en su término amino. Si el péptido contiene más de un grupo amino libre,
55 se pueden formar grandes complejos multiméricos.

En un aspecto, se pueden fusionar los péptidos o polipéptidos de la presente invención (por ejemplo, mediante clonación en marco) o ligar (por ejemplo, mediante acoplamiento químico) a inhibidores celulares tales como agentes antimicrobianos. Entre estos se incluyen los péptidos antimicrobianos, por ejemplo proteína
60 bactericida/potenciadora de la permeabilidad, proteínas antimicrobianas catiónicas, lisozimas, lactoferrinas y catelicidinas (por ejemplo, de neutrófilos; ver, por ejemplo, Hancock y Chapple, 1999, *Antimicrob. Agents Chemother.* 43:1317-1323; Ganz y Lehrer, 1997, *Curr. Opin. Hematol.* 4:53-58; Hancock et al., 1995, *Adv. Microb. Physiol.* 37:135-175). Los péptidos antimicrobianos incluyen además defensinas (por ejemplo, de células epiteliales

o neutrófilos) y proteínasa microbicidas de plaquetas (ver, por ejemplo, Hancock y Chapple, 1999, *Antimicrob. Agents Chemother.* 43:1317–1323). Entre los péptidos antimicrobianos adicionales se incluyen, aunque no a modo de limitación, gramicidina S, bacitracina, polimixina B, taquiplesina, batenecina (por ejemplo, batenecina de vacuno), ranalexina, cecropina A, indolicidina (por ejemplo, indolicidina de vacuno), y nisina (por ejemplo, nisina bacteriana).

También se incluyen como agentes antimicrobianos los ionóforos, que facilitan la transmisión de un ion (tal como sodio), a través de una barrera lipídica tal como una membrana celular. Dos compuestos de ionóforos especialmente adecuados para esta invención son el RUMENSINTM (Eli Lilly) y Lasalocid (Hoffman LaRoche). Entre otros ionóforos se incluyen, aunque no a modo de limitación, salinomicina, avoparcina, aridicina y actaplanina. Otros agentes antimicrobianos incluyen Monensin™ y azitromicina, metronidazol, streptomina, kanamicina y penicilina, como así también, en general, β-lactamas, aminoglucósidos, macrólidos, cloranfenicol, novobiocina, rifampina y fluoroquinolonas (ver, por ejemplo, Horn et al., 2003, *Applied Environ. Microbiol.* 69:74–83; Eckburg et al., 2003, *Infection Immunity* 71:591–596; Gijzen et al., 1991, *Applied Environ. Microbiol.* 57:1630–1634; Bonelo et al., 1984, *FEMS Microbiol. Lett.* 21:341–345; Huser et al., 1982, *Arch. Microbiol.* 132:1–9; Hilpert et al., 1981, *Zentbl. Bakteriologie Mikrobiol. Hyg. 1 Abt Orig. C* 2:21–31).

Los inhibidores particularmente ventajosos son compuestos que bloquean o interfieren con la metanogénesis, incluyendo ácido bromoetansulfónico, por ejemplo ácido 2-bromoetansulfónico (BES) o una sal del mismo, por ejemplo una sal de sodio. El molibdato de sodio (Mo) es un inhibidor de la reducción del sulfato y se lo puede utilizar con el ácido bromoetansulfónico. Entre otros compuestos antimetánogénicos se cuentan, aunque no a modo de limitación, nitrato, formato, metil fluoruro, cloroformo, hidrato de cloral, sulfito de sodio, etileno e hidrocarburos insaturados, acetileno, ácidos grasos tales como el ácido linoleico y ácido cis-oleico, ácidos grasos saturados tales como ácido behénico y ácido esteárico y también lumazina (por ejemplo, 2,4-pteridinediona). Otros compuestos incluyen 3-bromopropansulfonato (BPS), ácido propiónico y etil 2-butanoato.

También se incluyen como agentes antimicrobianos las enzimas líticas, incluyendo la lisozima, endolisina, lisozima, lisina, lisina de fagos, muralisina, muramidasa y virolisina. Las enzimas útiles exhiben la capacidad de hidrolizar enlaces específicos en la pared de la célula bacteriana. Las enzimas líticas específicas incluyen, aunque no a modo de limitación, glucosaminidasas, que hidrolizan los enlaces glucosídicos entre los amino azúcares (por ejemplo, el ácido N-acetilmurámico y N-acetilglucosamina) del péptidoglicano, amidasas, que escinden el ligadura N-acetilmuramoyl-L-alanina amida entre la hebra de glicano y el péptido entrecruzador, y endopeptidasas, que hidrolizan el ligadura entre péptidos (por ejemplo cisteína endopeptidasas) y las endoisopeptidasas que atacan la pseudomureína de los metanógenos de la familia Methanobacteriaceae.

Además, se incluyen los PNAs como agentes antimicrobianos. Los PNAs son híbridos péptido-ácido nucleico en los cuales se ha reemplazado la columna principal de fosfato por una columna principal acquiral y neutra preparada a partir de unidades de N-(2-aminoetil)-glicina (ver, por ejemplo, Eureka Bioscience Collection. PNA and Oligonucleotide Inhibitors of Human Telomerase. G. Gavory y S. Balasubramanian, Landes Bioscience, 2003). Las bases A, G, T, C se anexan al nitrógeno de amino de la columna vertebral por medio de ligaduras metilencarbonilo (P.E. Nielsen et al., *Science* 1991. 254: 1497–1500; M. Egholm et al., *Nature* 1993. 365: 566–568). Los PNAs enlazan secuencias complementarias con alta especificidad, y una mayor afinidad con respecto al ADN o ARN análogo (M. Egholm et al., supra). Los híbridos PNA/DNA o PNA/RNA exhiben además una mayor termoestabilidad que los correspondientes dúplexes DNA/DNA o DNA/RNA (M. Egholm et al., supra). Los PNAs también poseen alta estabilidad química y biológica, debido a la columna principal amídica no natural que no es reconocida por las nucleasas ni las proteasas (V. Demidov et al., *Biochem Pharmacol* 1994. 48: 1310–1313). Por lo general, los PNAs tienen una longitud de por lo menos 5 bases e incluyen una lisina terminal. Los PNA pueden ser pegilados para extender aun más su vida útil (Nielsen, P. E. et al. (1993) *Anticancer Drug Des.* 8:53–63).

En un aspecto específico, los péptidos o polipéptidos de la presente invención pueden estar fusionados (por ejemplo, por clonación en marco) o ligados (por ejemplo, mediante acoplamiento químico) a inhibidores celulares tales como anticuerpos o fragmentos de los mismos. Los anticuerpos o fragmentos de anticuerpos se pueden dirigir a las células microbianas o, especialmente, a las células metanógenas o a uno o más componentes celulares. Por ejemplo, se pueden tomar como blancos proteínas superficiales de las células, por ejemplo receptores. Se incluyen las moléculas de inmunoglobulina y porciones inmunológicamente activas de las moléculas de inmunoglobulina (Ig), es decir, moléculas que contienen un sitio de unión al antígeno que se une específicamente (inmunorreacciona con) un antígeno.

Los péptidos o polipéptidos de la presente invención son de uso específico en la selección como blanco de una célula microbiana, especialmente una célula metanógena. En ciertos aspectos, se pueden utilizar los péptidos y polipéptidos para fijarse a la pared o la membrana celular y/o permeabilizar la célula. Por tal motivo, los péptidos o polipéptidos se pueden utilizar para la unión transitoria o extendida a la célula o para penetrar en la pared o membrana celular y/o acumularse en el ambiente celular. Se entiende que los péptidos, polipéptidos, como así los

correspondientes polinucleótidos, vectores de expresión, células huésped y anticuerpos de acuerdo con la presente invención, se pueden utilizar para tomar como blanco diversos microbios, por ejemplo *Methanobrevibacter ruminantium*, que es un metanógeno común en los rumiantes, y *Methanobrevibacter smithii*, que es un metanógeno común en los humanos. Para realizar la selección como blanco, se puede poner en contacto a la célula microbiana con el péptido señal o el polipéptido que comprende el péptido, aislado de una o más fuentes naturales o producido por vectores de expresión y/o células huésped o por química sintética o semisintética descrita en detalle en la presente. En los aspectos específicos, se administra el péptido o polipéptido a los sujetos en forma de composición descrita en este documento en forma detallada, por ejemplo por medio del uso de un dispositivo de liberación lenta en el caso de los rumiantes.

10

En ciertas modalidades, el polipéptido se fusiona o liga a un inhibidor celular, por ejemplo un compuesto antimetanoagénico (por ejemplo, ácido bromoetansulfónico), un anticuerpo o fragmento de anticuerpo, enzima lítica, ácido nucleico peptídico, péptido antimicrobiano u otro antibiótico. El inhibidor peptídico o inhibidor polipeptídico es administrado a los sujetos en forma de composición para inhibir el desarrollo o replicación de las células microbianas, especialmente células metanógenas. La composición comprende, por ejemplo: a) un péptido señal aislado o un polipéptido que comprende este péptido o una alteración, fragmento, variante o derivado del mismo; b) un polinucleótido aislado o una alteración, fragmento, variante o derivado del mismo; c) un vector de expresión que comprende este polinucleótido o d) una célula huésped que comprende este vector de expresión. Las composiciones de la presente invención pueden ser envasadas específicamente como parte de kits para la selección como blanco, permeabilización y/o inhibición de células microbianas, especialmente células metanógenas, de acuerdo con los métodos descritos. Los kits comprenden por lo menos una composición de acuerdo con lo expuesto en la presente e instrucciones para su uso en la permeabilización de las células o para inhibir el desarrollo o replicación de metanógenos u otros microbios.

Como modalidad adicional, la invención se relaciona con una composición farmacéutica en combinación con un vehículo aceptable para uso farmacéutico, para usar con cualquiera de los métodos descritos anteriormente. Dichas composiciones farmacéuticas pueden comprender un péptido señal o un polipéptido que comprende este péptido, en combinación con un inhibidor celular. Por otro lado, las composiciones farmacéuticas pueden comprender un vector de expresión o célula huésped descrita en la presente. Las composiciones pueden ser administradas solas o en combinación con por lo menos un agente adicional, como por ejemplo un compuesto estabilizador, que se puede administrar en cualquier vehículo farmacéutico estéril biocompatible, incluyendo, aunque no a modo de limitación, solución salina, salina tamponada, dextrosa y agua. Las composiciones se pueden administrar a un sujeto solas o en combinación con otros agentes, fármacos (por ejemplo fármacos antimicrobianos) u hormonas.

Además de los ingredientes activos, estas composiciones farmacéuticas pueden contener vehículos aceptables para uso farmacéutico que comprenden excipientes y auxiliares que faciliten el procesamiento de los compuestos activos en preparaciones que se pueden utilizar en forma farmacéutica. Se pueden encontrar más detalles sobre técnicas de formulación y administración en la última edición de Remington's Pharmaceutical Sciences (Maack Publishing Co., Easton, PA). Las composiciones farmacéuticas utilizadas en la presente invención se pueden administrar por cualquier número de vías incluyendo, aunque no a modo de limitación, la vía oral, endovenosa, intramuscular, intraarterial, intramedular, intratecal, intraventricular, transdérmica, subcutánea, intraperitoneal, intranasal, enteral, tópica, sublingual o rectal.

Las composiciones farmacéuticas para la administración oral se pueden formular utilizando vehículos aceptables para uso farmacéutico muy conocidos en la técnica en dosis adecuadas para la administración oral. Dichos vehículos permiten la formulación de las composiciones farmacéuticas en forma de comprimidos, píldoras, grageas, cápsulas, líquidos, geles, jarabes, geles, suspensiones y demás, para ser ingeridos por el sujeto. Las composiciones farmacéuticas para uso oral se pueden obtener mediante la combinación de los compuestos activos con un excipiente sólido, optativamente la molienda de la mezcla obtenida y el procesamiento de la mezcla de gránulos, después de agregar los auxiliares adecuados, si resulta conveniente, para obtener núcleos de comprimidos o grageas. Los excipientes adecuados son materiales de carga de carbohidratos o proteínas tales como azúcares, incluyendo lactosa, sacarosa, manitol o sorbitol; almidón de maíz, trigo, arroz, papa u otras plantas; celulosa, como por ejemplo metil celulosa, hidroxipropilmetil-celulosa o carboximetilcelulosa sódica; gomas, incluyendo arábica y tragacanto y proteínas tales como gelatina y colágeno. Si conviene, se pueden agregar agentes desintegrantes o solubilizantes tales como polivinil pirrolidona entrecruzada, agar, ácido alginico o una sal del mismo tal como alginato de sodio.

Las preparaciones farmacéuticas que se pueden utilizar por vía oral incluyen las cápsulas de calce a presión hechas de gelatina, así como las cápsulas blandas selladas hechas de gelatina y un revestimiento tal como glicerol o sorbitol. Las cápsulas calzadas a presión pueden contener ingredientes activos mezclados con un material de carga o aglutinantes tales como lactosa o almidones, lubricantes tales como talco o estearato de magnesio y, optativamente, estabilizadores. En las cápsulas blandas, los compuestos activos pueden estar disueltos o suspendidos en líquidos adecuados, tales como aceites grasos, líquido o polietilenglicol líquido con o sin estabilizadores. Los núcleos de grageas se pueden utilizar en combinación con revestimientos adecuados tales

como soluciones azucaradas concentradas, que pueden contener además goma arábica, talco, polivinilpirrolidona, gel de carbopol, polietilenglicol y/o dióxido de titanio, soluciones de laca y solventes o mezclas de solventes orgánicos adecuados. Se pueden agregar tinturas o pigmentos a los revestimientos de los comprimidos o grageas para la identificación del producto o a fin de caracterizar la cantidad del compuesto activo, es decir, la dosis.

5

Las formulaciones farmacéuticas adecuadas para la administración parenteral pueden ser formuladas en soluciones acuosas, preferentemente en amortiguadores fisiológicamente compatibles tales como solución de Hanks, solución de Ringer o salina con amortiguador fisiológico. Las suspensiones acuosas para inyección pueden contener sustancias que aumenten la viscosidad de la suspensión, tales como carboximetil celulosa sódica, sorbitol o dextrano. Además, se pueden preparar suspensiones de los compuestos activos en forma de suspensiones oleosas para inyección apropiadas. Los solventes lipofílicos apropiados incluyen aceites grasos tales como aceite de sésamo o ésteres sintéticos de ácidos grasos tales como oleato de etilo o triglicéridos o liposomas. También se pueden utilizar amino polímeros policatiónicos no lipídicos para la administración. Optativamente, la suspensión puede contener además estabilizadores o agentes adecuados que aumenten la solubilidad de los compuestos para permitir la preparación de soluciones altamente concentradas. Para la administración tópica o nasal, se utilizan en la formulación agentes de penetración apropiados para la barrera específica que se ha de permear. Dichos agentes de penetración son generalmente conocidos en la técnica.

Las composiciones farmacéuticas de la presente invención se pueden elaborar de manera conocida en el medio, por ejemplo por medio de procesos convencionales de mezclado, disolución, granulación, preparación de grageas, levigación, emulsionamiento, encapsulación, atrapamiento o liofilización. La composición farmacéutica puede presentarse en forma de sal y puede estar conformada con muchos ácidos, entre los que se incluyen, aunque no a modo de limitación, los ácidos clorhídrico, sulfúrico, acético, láctico, tartárico, málico, succínico, etc. Las sales tienden a ser más solubles en solventes acuosos o protónicos de otro tipo que las correspondientes formas de base libre. En otros casos, la preparación preferida puede consistir en un polvo liofilizado que puede contener cualquiera o todos los siguientes: histidina 1–50 mM, 0.1%–2% de sacarosa y 2–7% de manitol, a un pH en el rango de 4,5 a 5,5, que se combinan con un amortiguador antes de usar. Una vez preparadas las composiciones farmacéuticas, se las puede colocar en un recipiente apropiado y rotular para el tratamiento de un trastorno adecuado. Para la administración de una composición de acuerdo con la presente invención, dicha rotulación incluiría la cantidad, frecuencia y método de administración.

Las composiciones farmacéuticas adecuadas para usar en la invención incluyen composiciones en las cuales los ingredientes activos están contenidos en una cantidad efectiva para obtener el fin pretendido. Para cualquier compuesto, se puede estimar la dosis terapéuticamente efectiva en ensayos celulares, por ejemplo en células microbianas o, en particular, células metanógenas o en modelos animales, habitualmente ratones, conejos, perros o cerdos, o en especies de rumiantes tales como ovejas, vacas, ciervos y cabras. También se puede utilizar el modelo animal para determinar la concentración y la vía de administración apropiadas. Seguidamente se puede utilizar esa información para determinar las dosis y vías de administración útiles en un sujeto. Las cantidades de dosificación normal pueden variar entre 0.1 y 100.000 microgramos, hasta una dosis total de aproximadamente 1 g, dependiendo de la vía de administración. Se presentan lineamientos en cuanto a las dosis específicas y los métodos de administración en la literatura y generalmente están a disposición de los profesionales en la técnica. Los expertos en la técnica han de emplear diferentes formulaciones para los polinucleótidos que en el caso de los péptidos o polipéptidos. Del mismo modo, la administración de polinucleótidos o polipéptidos han de ser específicas para las células, condiciones, ubicaciones particulares, etc.

45

Las terapias basadas en péptidos y polipéptidos son muy conocidas y los métodos de elaboración de dichas composiciones ya han sido cimentados en la técnica. Se han descrito ejemplos de terapias con péptidos y polipéptidos y su elaboración, con respecto, por ejemplo, adenileuquin difitox, octreótido, vapreótido, lanreótido, péptidos de la serie RC–3940, decapeptilo, lupron, zoladex, cetorelix (ver, por ejemplo, Lu et al., 2006, AAPS J 8:E466–472), hemocidinas, estafoaínas (ver, por ejemplo, Dubin et al., 2005, Acta Biochemica Polonica, 52:633–638), como así también indolicidina, defensinas, lantibióticos, microcidina B17, histatinas y maganina (ver, por ejemplo, Yeaman y Yount, 2003, Pharmacol Rev 55:27–55). También se puede encontrar una guía general para las terapias con péptidos y polipéptidos en el trabajo de Degim et al., 2007, Curr Pharm Des 13:99–117 y Shai et al., 2006, Curr Prot Pept Sci, 7:479–486. Los fármacos basados en péptidos recientemente apropiados incluyen Hematide™ (agente estimulante de la eritropoyesis basado en péptidos, Affymax, Inc.), Exenatida (exendina–4 sintética, Amylin/Eli Lilly), Natreacor (nesiritida, péptido natriurético, Scios), Plenaxis (abarelix, Praecis Pharmaceuticals) y SecreFlo (secretina, Repligen).

La dosis exacta debe ser determinada por el profesional, tomando en cuenta los factores relacionados con el sujeto que requiere tratamiento. La dosis y la administración se ajustan de manera que otorguen niveles suficientes del agente activo o para mantener el efecto deseado. Los factores que se pueden tomar en cuenta incluyen la severidad de la patología, el estado de salud general del sujeto, la edad, peso y sexo del sujeto, la dieta, tiempo y frecuencia de administración, la combinación (o combinaciones) de fármacos, las sensibilidad a la reacción y la

60

tolerancia/respuesta a la terapia. Se pueden administrar composiciones farmacéuticas de acción prolongada cada 3 a 4 días, cada semana o una vez cada dos semanas, dependiendo de la vida media y la velocidad de clearance de la formulación específica.

- 5 Son de particular utilidad para las composiciones de la presente invención (por ejemplo, composiciones farmacéuticas) las fórmulas o mecanismos de liberación lenta. Por ejemplo, los dispositivos intra-ruminales incluyen, aunque no a modo de limitación, el rango de bolo de Time Capsule™ Bolus de Agri-Feeds Ltd., Nueva Zelanda, desarrollada originariamente en AgResearch de Nufarm Health & Sciences, división de Nufarm Ltd., Auckland, Nueva Zelanda, descrita en AU 35908178, PCT/AU81/100082, y Laby et al., 1984, Can. J. Anim. Sci. 64 (Suppl.),
10 337-8. Como ejemplo específico, el dispositivo puede incluir un resorte y émbolo que empujan a la composición contra un orificio en el extremo de un barril.

Como modalidad adicional, la invención se relaciona con una composición para un suplemento acuoso, por ejemplo remojando la composición o suplemento alimentario, por ejemplo un componente de pienso para rumiantes, para
15 usar con cualquiera de los métodos antes descritos. En aspectos específicos, el suplemento alimentario comprende por lo menos un material vegetal comestible y un péptido o polipéptido de la presente invención. Por otro lado, el suplemento alimentario comprende por lo menos un material vegetal comestible y un péptido o polipéptido o polinucleótido que codifica un péptido o polipéptido descrito en la presente, por ejemplo como vector de expresión o célula huésped que comprende el vector de expresión. Específicamente, la composición comprende además un
20 inhibidor celular, fusionado o ligado a la secuencia obtenida. El material vegetal preferido incluye cualquiera de los siguientes: heno, pasto, grano o harina, por ejemplo heno de legumbres, heno de hierbas, maíz para ensilaje, pastos para ensilaje, legumbres para ensilaje, granos de maíz, avena, cebada, granos para destiladores, granos para cerveceros, harina de soja y harina de semillas de algodón. Específicamente, los pastos para ensilaje son útiles como composición alimenticia para rumiantes. El material de plantas puede ser genéticamente modificado de
25 manera que tenga uno o más componentes de la presente invención, por ejemplo uno o más polipéptidos o péptidos, polinucleótidos o vectores. En otra modalidad, se pueden utilizar anticuerpos que se unen específicamente a los péptidos, polipéptidos o polinucleótidos de la presente invención para determinar la presencia de microbios, especialmente metanógenos, o en ensayos para monitorear los niveles de esos microbios. Los anticuerpos ventajosos para fines de diagnóstico se pueden preparar de la misma manera que los descritos anteriormente. Los
30 ensayos de diagnóstico incluyen métodos que utilizan el anticuerpo y un marcador para detectar un péptido o polipéptido en fluidos o extractos de células o tejidos del cuerpo humano. Los anticuerpos se pueden utilizar con o sin modificación y se los puede marcar uniéndolos, ya sea en forma covalente o no covalente, con una molécula reportera. Se puede utilizar una amplia variedad de moléculas reporteras que son conocidas en la técnica, varias de las cuales fueron descritas anteriormente.

35 En la técnica se conoce una variedad de protocolos para medir los niveles de un péptido, polipéptido o polinucleótido (por ejemplo, ELISA, RIA, FACS, y transferencias), y otorgan una base para determinar la presencia o los niveles de un microbio, especialmente un metanógeno. Se establecen los niveles normales o standard combinando los fluidos corporales o los extractos celulares extraídos de sujetos normales (por ejemplo humanos o rumiantes normales, con
40 el anticuerpo en condiciones adecuadas para la formación de complejos. La cantidad de formación de complejos standard se puede cuantificar por diversos métodos, aunque preferentemente por medios fotométricos. Se comparan las cantidades de péptido, polipéptido o polinucleótido expresadas en muestras de un sujeto, control y tratadas (por ejemplo, muestras de sujetos tratados) con los valores standard. La desviación entre los valores standard y del sujeto establece los parámetros para determinar la presencia o los niveles del microbio.

45 En una modalidad específica de la presente invención, los polinucleótidos se pueden utilizar con fines de diagnóstico utilizando técnicas de hibridación y/o amplificación específicas. Los polinucleótidos que se pueden utilizar incluyen oligonucleótidos, moléculas de ARN y ADN complementario y PNA's. Los polinucleótidos se pueden utilizar para detectar y cuantificar la expresión génica en las muestras en las cuales se puede correlacionar la expresión con la
50 presencia o niveles de un microbio. Se puede emplear el ensayo de diagnóstico para distinguir entre la ausencia, presencia y modificación de los niveles de microbios, y para monitorear los niveles durante la intervención terapéutica.

En un aspecto, se puede utilizar la hibridación con sondas de PCR para identificar secuencias de ácido nucleico, especialmente secuencias genómicas, que codifican los péptidos o polipéptidos de la presente invención. La
55 especificidad de la sonda, ya sea que esté preparada a partir de una región altamente específica, por ejemplo 10 nucleótidos individuales de la región reguladora 5' o una región menos específica, por ejemplo en la región codificadora 3', y la rigurosidad de la hibridación o amplificación (máxima, alta, intermedia o baja) determinará si la sonda identifica sólo secuencias de origen natural, alelos o secuencias relacionadas. También se pueden utilizar
60 sondas para la detección de secuencias relacionadas, y deben contener preferentemente por lo menos 50 % de los nucleótidos de cualquiera de las secuencias codificadoras. Las sondas de hibridación de la presente invención pueden ser de ADN o ARN y derivar de la secuencia de nucleótidos de SEC. ID. NO: 173-341 o 342-533 o

complementos o secuencias modificadas de las mismas o de secuencias genómicas que incluyen promotores, elementos potenciadores e intrones de la secuencia de origen natural.

- Los medios para la producción de sondas de hibridación específica para los ADN incluyen la clonación de
- 5 secuencias de ácidos nucleicos en vectores para la producción de sondas de ARNm. Dichos vectores son conocidos en la técnica, existen en el comercio y se los puede utilizar para sintetizar sondas de ARN in vitro por medio de la adición de las ARN polimerasa apropiadas y los nucleótidos marcados apropiados. Las sondas de hibridación pueden ser marcadas por una variedad de grupos reporteros, por ejemplo radionúclidos tales como ³²P o ³⁵S o
- 10 marcas enzimáticas, tales como fosfatasa alcalina acoplada a la sonda por medio de sistemas de acoplamiento de avidina/biotina y demás. Los polinucleótidos se pueden utilizar en análisis Southern o northern, de transferencias de puntos u otras tecnologías basadas en membranas, en tecnologías de PCR o en ensayos de varilla indicadora, perno, ELISA o micromatrices que utilizan fluidos o tejidos de biopsias de sujetos para detectar la presencia o los niveles de un microbio. Dichos métodos cualitativos o cuantitativos son muy conocidos en la técnica.
- 15 En un aspecto específico, las secuencias de ácido nucleico pueden ser de utilidad en diversos ensayos marcados por métodos standard y agregados a una muestra de fluido o tejido de un sujeto en condiciones adecuadas para la hibridación y/o amplificación. Después de un período de incubación adecuado, se lava la mezcla y se cuantifica la señal y se l^o compara con un valor standard. Si la cantidad de señal presente en la muestra de ensayo está significativamente modificada con respecto a la de una muestra control comparable, la presencia de niveles
- 20 alterados de secuencias nucleotídicas en la muestra indica la presencia o niveles del microbio. Dichos ensayos se pueden utilizar asimismo para evaluar la eficacia de un régimen de tratamiento específico en estudios con animales, en ensayos clínicos o en el monitoreo del tratamiento impartido a un sujeto.

- Para otorgar una base para el diagnóstico de la presencia o niveles de un microbio, se establece un perfil normal o
- 25 standard para la expresión. Esto se puede lograr combinando fluidos corporales o extractos celulares tomados de sujetos normales con un polinucleótido o un fragmento del mismo, en condiciones adecuadas para la hibridación y/o amplificación. Se pueden cuantificar los niveles standard comparando los valores obtenidos de sujetos normales con los obtenidos en un experimento en que se utiliza una cantidad conocida de un polinucleótido sustancialmente purificado. Se pueden comparar los valores standard obtenidos de las muestras normales con los valores obtenidos
- 30 de muestras de sujetos tratados para el desarrollo de microbios. Se utiliza la desviación entre los valores standard y los del sujeto para establecer la presencia o los niveles del microbio.

- Una vez identificado el microbio e iniciado un protocolo de tratamiento, se pueden repetir los ensayos de hibridación y/o amplificación en forma regular para evaluar si el nivel de expresión en el sujeto comienza a decrecer con
- 35 respecto al observado en el sujeto normal. Se pueden utilizar los resultados obtenidos de ensayos sucesivos para demostrar la eficacia del tratamiento en el curso de un período en el rango de varios días a meses.

- Los usos de diagnóstico específicos para los oligonucleótidos diseñados a partir de secuencias de ácidos nucleicos pueden implicar el uso de PCR. Dichos oligómeros pueden ser sintetizados químicamente, generados
- 40 enzimáticamente o producidos in vitro. Los oligómeros consisten preferentemente en dos secuencias nucleotídicas, una con orientación en sentido directo (5'.fwdarw.3') y otra con orientación antisentido (3'.fwdarw.5'), empleadas en condiciones optimizadas para la identificación de un gen o condición específica. Se pueden utilizar dos oligómeros iguales, series anidadas de oligómeros o incluso un grupo degenerado de oligómeros en condiciones menos rigurosas para la detección y/o cuantificación de secuencias de ADN o ARN estrechamente relacionadas.

- 45 Los métodos que también se pueden utilizar para cuantificar la expresión incluyen nucleótidos para radiomarcación o biotinilación, coamplificación de un ácido nucleico control y curvas standard sobre las cuales se interpolan los resultados experimentales (Melby, P. C. et al. (1993) *J. Immunol. Methods*, 159:235–244; Duplaa, C. et al. (1993) *Anal. Biochem.* 229–236). Se puede acelerar la velocidad de cuantificación de múltiples muestras corriendo el ensayo en un formato ELISA en el cual se presenta el oligómero de interés en diversas diluciones y una respuesta
- 50 espectrofotométrica o colorimétrica da una cuantificación rápida.

- En otras modalidades, se pueden utilizar oligonucleótidos o fragmentos de mayor longitud derivados de cualquiera de los polinucleótidos descritos en la presente como blancos en una micromatriz. Se puede utilizar la micromatriz para monitorear el nivel de expresión de grandes números de genes simultáneamente (para producir una imagen
- 55 transcripta) y para identificar las variantes genéticas, mutaciones y polimorfismos. Se puede utilizar esta información para determinar la función génica, para comprender la base genética de la enfermedad, para diagnosticar la enfermedad y para desarrollar y monitorear las actividades de los agentes terapéuticos. En una modalidad, se prepara la micromatriz y se la utiliza de acuerdo con métodos conocidos en la técnica tales como los descritos en la solicitud PCT WO 95/11995 (Chee et al.), Lockhart, D. J. et al. (1996; *Nat. Biotech.* 14: 1675–1680) y Schena, M. et
- 60 al. (1996; *Proc. Natl. Acad. Sci.* 93: 10614–10619).

En un aspecto, los oligonucleótidos pueden ser sintetizados sobre la superficie de una micromatriz utilizando un procedimiento de acoplamiento químico y un aparato de aplicación de chorro de tinta tal como el descrito en la

- solicitud PCT WO95/251116 (Balduchweiler et al.). En otro aspecto, se puede utilizar una matriz "de parrilla" análoga a una transferencia de puntos o ranura (aparato HYBRIDOT, Life Technologies) para disponer y lugar fragmentos de ADNc u oligonucleótidos a la superficie de un sustrato utilizando un sistema de vacío, procedimientos de enlace térmicos, por UV, mecánicos o químicos. En un aspecto más, se puede producir una matriz a mano o
 5 utilizando dispositivos, materiales y máquinas existentes (incluyendo pipeteadoras multicanales o instrumentos robóticos; Brinkmann, Westbury, N.Y.) y pueden contener 8, 24, 96, 384, 1536 o 6144 oligonucleótidos o cualquier otro múltiplo de 2 a 1.000.000 que se preste al uso eficiente de los instrumentos existentes en el comercio. Para llevar a cabo el análisis de las muestras empleando las micromatrices, se extraen polinucleótidos de una muestra biológica. Se pueden obtener las muestras biológicas de cualquier fluido corporal (sangre, orina, saliva,
 10 flema, jugos gástricos, etc.), células cultivadas, biopsias u otras preparaciones de tejidos. Para producir sondas, se utilizan los polinucleótidos extraídos de la muestra para producir secuencias de ácidos nucleicos que sean complementarias de los ácidos nucleicos dispuestos en la micromatriz. Si la micromatriz consiste en ADNc, son apropiadas las sondas de ARNm. Por lo tanto, en un aspecto, se utiliza ARNm para producir ADNc que, a su vez, y en presencia de nucleótidos fluorescentes, para producir fragmentos o sondas de ARN antisentido. Estas sondas
 15 con marca fluorescente son incubadas con la micromatriz, de manera que las secuencias de la sonda se hibriden a los oligonucleótidos de ADNc de la micromatriz. En otro aspecto, las secuencias de ácido nucleico utilizadas como sondas pueden incluir polinucleótidos, fragmentos y secuencias complementarias o antisentido utilizando enzimas de restricción, tecnologías de PCR y kits de oligomarcación (Amersham Pharmacia Biotech) muy conocidos en el campo de la tecnología de hibridación.
 20
- En otra modalidad de la presente invención, se pueden utilizar los péptidos o polipéptidos de la presente invención o fragmentos funcionales o inmunogénicos de los mismos, para clasificar bibliotecas de compuestos en cualquiera de una variedad de técnicas de clasificación de fármacos. El fragmento empleado en esa clasificación puede estar libre en solución, fijo a un soporte sólido, soportado sobre la superficie de una célula o ubicado en el interior de la célula.
 25 Se puede medir la formación de complejos de unión entre el péptido o polipéptido y el agente que se está analizando.

- Una técnica para la clasificación de fármacos que se puede utilizar da lugar a la clasificación de alto rendimiento de compuestos con afinidad de unión con el péptido o polipéptido de interés, de acuerdo con lo descrito en la solicitud
 30 PCT publicada WO 84/03564. En este método, se sintetizan grandes números de pequeños compuestos de ensayo diferentes sobre un sustrato sólido, como por ejemplo pernos plásticos o alguna otra superficie. Se hace reaccionar los compuestos de ensayo con el péptido o polipéptido o fragmentos del mismo y se los lava. A continuación se detecta el péptido o polipéptido unido mediante métodos muy conocidos en la técnica. También se puede aplicar el péptido o polipéptido purificado directamente sobre placas para usar en las técnicas de clasificación de fármacos
 35 antes mencionadas. Por otro lado, se pueden utilizar anticuerpos no neutralizantes para capturar el péptido e inmovilizarlo sobre un soporte sólido.

- En otra técnica, se pueden utilizar ensayos de clasificación competitiva de fármacos en los cuales anticuerpos con capacidad para unirse al péptido o polipéptido específicamente compiten con un compuesto de ensayo por la unión
 40 al péptido o polipéptido. De esta manera, se pueden utilizar los anticuerpos para detectar la presencia de un compuesto de ensayo que comparte uno o más sitios de unión al antígeno con el anticuerpo.

EJEMPLOS

- 45 Los ejemplos descritos en la presente se presentan con el propósito de ilustrar modalidades de la invención. Otras modalidades, métodos y tipos de análisis están dentro del alcance de Los expertos en la técnica del diagnóstico molecular y no es necesario describirlas en forma detallada en este documento. Otras modalidades dentro del alcance de la técnica se consideran partes de la presente invención.

50 EJEMPLO 1: Materiales y métodos

Estimado del tamaño del genoma

- Se cultivó la cepa M1T de *Methanobrevibacter ruminantium* (DSM1093) en medio BY+ (medio basal, Joblin et al.,
 55 1990) que consiste en [g/l]: NaCl (1), KH₂PO₄ (0,5), (NH₄)₂SO₄ (0,25), CaCl₂·2H₂O (0,13), MgSO₄·7H₂O (0,2), K₂HPO₄ (1), fluido de rumen clarificado (300 ml) dH₂O (360 ml), NaHCO₃ (5), resazurina (0,2 ml) L-cisteína-HCl (0,5), extracto de levadura (2), y solución de oligoelementos de Balch (10 ml) (elementos trazo agregados; Balch et al., 1979) que consiste en: (g/l) ácido nitriloacético (1,5), MgSO₄·7H₂O (3), MnSO₄·H₂O (0,5), NaCl (1), FeSO₄·7H₂O (0,1), CoCl₂·6H₂O (0,1), CaCl₂ (0,1), ZnSO₄·7H₂O (0,1), CuSO₄·5H₂O (0,01), AlK(SO₄)₂·12H₂O
 60 (0,01), H₃BO₃ (0,01), Na₂MoO₄·2H₂O (0,01), NiSO₄·6H₂O (0,03), Na₂SeO₃ (0,02) y Na₂WO₄·2H₂O (0,02). Se extrajo ADN genómico utilizando un método de congelación y molienda. Se cosecharon las células por centrifugación y se colocó el pellet celular en un mortero previamente enfriado, congelado con nitrógeno líquido y suavemente molido hasta obtener un polvo fino utilizando un mortero con mano esterilizado previamente enfriado.

Se embutieron los homogenatos celulares en tapones de agarosa y se llevaron a cabo sucesivas manipulaciones en los tapones para reducir el corte físico de ADN genómico. Se llevaron a cabo digestiones con endonucleasas de restricción y se separaron los fragmentos de ADN utilizando electroforesis en gel de campo pulsado (PFGE).

Clonación y secuenciación de ADN

5

El ADN del genoma de *M. ruminantium* fue secuenciado por Agencourt Biosciences Corporation (Massachusetts, Estados Unidos) utilizando una técnica de clonación con pistola aleatoria (Fleischmann et al., 1995) y por MacroGen Corporation (Rockville, MD, Estados Unidos) usando pirosecuenciación. Se construyeron bibliotecas de ADN de *M. ruminantium* en *Escherichia coli* mediante alteración física aleatoria del ADN genómico y separación de los fragmentos por electroforesis en gel. Se recuperaron grandes fragmentos en el rango de 40 Kb del gel y se los utilizó para generar una biblioteca de insertos de fásmidos de gran tamaño. Se recuperaron los fragmentos de ADN en el rango de 2 y 4 kb y se los utilizó para generar una biblioteca de insertos plasmídicos de tamaño pequeño. Se desarrollaron los clones producidos como resultado de las bibliotecas de grandes y pequeños insertos, se recuperó su ADN de fásmidos o plásmidos y se lo secuenció utilizando tecnología de secuenciación de alto rendimiento. Se secuenciaron suficientes clones para dar, teóricamente, una cobertura de 8 veces el genoma de *M. ruminantium*. Se realizó la pirosecuenciación en los fragmentos de ADN genómico cortados al azar para dar una cobertura teórica final de 10 veces.

Ensamble y análisis de Secuencias

20

Se alinearon las secuencias de ADN para encontrar traslajos y se las ensambló para formar secuencia contigua (contig) utilizando Paracel Genome Assembler (Paracel Inc, CA, USA) y el paquete Staden (Staden et al., 1998) en combinación con secuencias tanto de PCR normal como inversa. Se analizaron las contiguas utilizando el localizador de marcos de lectura abiertos (ORF) GLIMMER (Gene Locator Interpolated Markov Model ER, Salzberg et al., 1998) y se analizó cada ORF por BLAST (Basic Local Alignment Search Tool (Altschul et al., 1997) contra bases de datos de nucleótidos no redundantes y proteínas del National Center for Biotechnology Information (NCBI).

25

Se unieron las contiguas de la secuencia de la fase preliminar de 8 veces al azar por medio del ligadura artificial de las secuencias para generar una "pseudomolécula" y se la sometió al Institute for Genomic Research (TIGR, DC, Estados Unidos) para la autoanotación. Se ensamblaron los contiguos de la pirosecuenciación de 10 veces utilizando GLIMMER y se autoanotaron los ORFs empleando GAMOLA (Global Annotation of Multiplexed On-site Blasted DNA sequences; Altermann and Klaenhammer, 2003). Seguidamente se verificaron manualmente las anotaciones automáticas. Se categorizaron los ORF de acuerdo con su función utilizando la base de datos de cúmulos de proteínas ortólogas (COG) (umbral $1e-02$) (Tatusov et al., 2001).

30

35

Se determinaron los motivos proteicos por HMMER (protocolo de transferencia de hipertexto://hmmer.wustl.edu) utilizando bibliotecas PFAM HMM y TIGRFAM, con alineación global y local (protocolo de transferencia de hipertexto://pfam.wustl.edu) y los modelos standard y en modalidad de fragmentos TIGRFAM HMMs (protocolo de transferencia de hipertexto://www.tigr.org/TIGRFAMs) respectivamente (umbral $1e-02$). Se identificaron los ARNt utilizando TRNASCAN-SE (Lowe y Eddy, 1997) y se identificaron repeticiones nucleotídicas empleando el paquete de software KODON (Applied Maths, Austin, TX, Estados Unidos) y REPUTER (Kurtz y Schleiermacher, 1999). Se construyeron visualizaciones del atlas del genoma utilizando GENEWIZ (Jensen et al., 1999). Se llevaron a cabo reconstrucciones de las vías del ORFeoma previsto de *M. ruminantium* en combinación con la base de datos en línea KEGG (Kyoto Encyclopedia of Genes and Genomes, Kanehisa et al., 2004) utilizando software desarrollado internamente (PathwayVoyager; Altermann y Klaenhammer, 2005).

40

45

Identificación de Péptidos Señal

Hasta la fecha no hay modelo de péptido señal para los archaea. Simplemente hay muy pocas proteínas secretorias verificadas experimentalmente disponibles para Archaea para generar un modelo específico. Por esta razón, se analizaron las secuencias de los marcos de lectura abiertos (ORF) para determinar la presencia de péptidos señal utilizando la Versión 3.0 de SignalP (Bendtsen et al., 2004) generados contra modelos gram positivos, Gram negativos y Eucariotas. Se utilizó SignalP-HMM (modelos ocultos de Markov) para discriminar entre el péptido señal y los ORF de péptidos no señal, en tanto que se utilizó SignalP-NN (redes neurales) para la predicción de los sitios de escisión de acuerdo con lo descrito por Emanuelsson et al., 2007.

50

55

SignalP predice la presencia y ubicación de los sitios de escisión de péptidos señal en las secuencias de aminoácidos de diferentes organismos. El método incorpora una predicción de los sitios de escisión y una predicción de péptidos señal/péptidos no señal basada en una combinación de varias redes neurales artificiales y modelos ocultos de Markov. Se alinearon las secuencias peptídicas señal identificadas de la serie de datos de Gram +ve y se calculó una secuencia consenso empleando el programa AlignX de Vector NTI (versión 9.1.0, Invitrogen Corporation). Se identificó el núcleo hidrófobo conservado por análisis de hidrofobicidad de los aminoácidos.

60

Se identificó una serie de datos de consenso de los tres modelos SignalP y se alinearon las secuencias de péptidos señal correspondientes empleando ClustalW (Larkin et al., 2007) y se las editó utilizando BioEdit (protocolo de transferencia de hipertexto://world wide web.mbio.ncsu.edu/BioEdit/bioedit.html). Se generó un logo de secuencia proteica (FIG. 3A) utilizando LogoBar (Perez-Bercoff et al., 2006) para representar la información presente en este

5 múltiple alineamiento de secuencias. En este estudio, se consideró que los ORF que contenían un péptido señal y tres o más dominios transmembrana eran proteínas de membrana y quedaron excluidos de los estudios posteriores. Se tomó el mejor puntaje Y de cada uno de los tres modelos como sitio de escisión putativo (FIG. 6). Se calculó el uso de codones optimizado (FIG. 9) utilizando un script perl interno opt_codons.pl (Altermann, E) basado en la tabla de codones de Escherichia coli K12.

10

Síntesis peptídica y marcación con fluoresceína

Se sintetizó el péptido consenso núcleo utilizando comercialmente el servicio de péptidos especializados de Invitrogen (Invitrogen NZ Ltd). Se sintetizó el péptido utilizando la química Fmoc en pequeña escala (10–12 mg) y se

15 lo purificó por HPLC a una pureza >95%. Se marcó el péptido en la lisina N-terminal (K) con fluoresceína.

Ensayo de Permeabilización de las células

Se siguió el ingreso del péptido marcado al interior de las células de *M. ruminantium* por el ensayo de fluorescencia.

20 Se desarrolló un cultivo de *M. ruminantium* en 10 ml de medio BY+ y se lo recogió por centrifugación a 10.000 x g por espacio de 10 min a 4°C. Las células fueron transferidas a un tubo Eppendorf de polipropileno de 1,5 ml y se las lavó en 1 ml de amortiguador TE (Tris-HCl 10 mM, ácido etilendiamino tetraacético 1 mM, pH 8) se las recogió por centrifugación en una microcentrífuga a 13.000 x g por espacio de 10 min a 4°C. Se resuspendió a las células (aproximadamente 1×10^8) en un volumen total de 200 μ l de amortiguador TE y se agregaron 20 μ g de péptido

25 marcado con fluoresceína. Se incubó la mezcla por espacio de 30 minutos a 37°C y luego se la centrifugó a 13.000 x g durante 10 min a 4°C. Se retuvo el fluido celular por encima del pellet de células que constituyó la fracción del sobrenadante. Se lavó tres veces el pellet celular con 200 μ l de amortiguador TE resuspendiendo repetidamente las células en amortiguador y centrifugándolas a 13.000 x g por espacio de 10 min a 4°C.

30 Se agruparon los lavados y se constituyó la fracción de lavado de las células. Se resuspendió el pellet celular que quedó después del tercer lavado en 200 μ l de amortiguador TE que contenía 1 % de dodecilsulfato de sodio. Se centrifugaron las células a 13.000 x g por espacio de 10 min a 4°C para formar pellets con las células y se recolectó el fluido por encima del pellet celular y se constituyó la fracción asociada a las células. Se congeló el pellet celular restante en N2 líquido y se sometió a las células a alteración física mediante la molienda el pellet congelado con una

35 varilla de vidrio. Se centrifugó el homogenato celular obtenido a 20.000 x g por espacio de 30 min a 4°C. Se colectó el fluido presente por encima de las células que representaba la fracción intracelular, en tanto se resuspendía el material restante reducido a pellets en amortiguador TE, representando la fracción de pared / membrana celular. Se midió la fluorescencia en cada una de estas fracciones sellando una muestra de cada fracción en un capilar de vidrio y midiendo la fluorescencia emitida a 510–533 nm contra standards de péptidos marcados con fluoresceína

40 utilizando el detector de fluorescencia (Canal 1) de un Lightcycler (Roche).

EJEMPLO 2: Resultados experimentales

El cálculo del tamaño del genoma de *M. ruminantium* mediante digestión con enzimas de restricción de ADN

45 genómico y la determinación del tamaño de los fragmentos por medio de PFGE, indicó un solo cromosoma de aproximadamente 2.5–2.9 Mb. La secuenciación inicial de clones insertados de gran y pequeño tamaño (cobertura proyectada de 6 veces) y el ensamble de la secuencia en contiguos indicó que una región de 40 Kb del genoma estaba sumamente sobre-representada (>20 veces), especialmente dentro de la biblioteca de pequeños insertos.

50 Debido a esta inclinación por las grandes secuencias, se llevó a cabo la secuenciación adicional (cobertura teórica del genoma de 2 veces) sólo correspondiente a clones de grandes insertos para dar una cobertura final de 8 veces producida por la secuenciación de Sanger. Se ensambló la fase proyectada de 8 veces formando 756 contiguos que estaban ligados por medio de 105 andamios. Se llevó a cabo otra pirosecuenciación hasta obtener una cobertura adicional de ~ 10 veces y la incorporación de estas secuencias al ensamble dio lugar a que el número de contiguos

55 cayera a 27. El subsiguiente cierre de las brechas utilizando técnicas de PCR inversa y de largo plazo redujo el número de contiguos a 14.

La longitud combinada de la secuencia de 14 contiguos indica que el genoma es ligeramente mayor (2.937.347 pb) que el tamaño estimado por PFGE (FIG. 1A) y significativamente mayor que su pariente más cercano, *M. smithii* (1,9

60 Mb). El % G+C de 32,64% se aproxima al rango informado de 27,5% a 31,6% dado a conocer respecto de las cepas de *M. ruminantium* (Balch et al, 1979). El análisis de la secuencia predice 2.239 ORFs y el número total de aciertos correspondientes a las familias proteicas (TIGRFam y PFam) y se dan a conocer los Cúmulos de Grupos Ortólogos (COGs) en la FIG. 1B. Todos los genes que se estimaban involucrados en la metanogénesis de H₂ + CO₂ y formiato

están presentes (FIG. 1C) Sin embargo, la secuencia proyectada de *M. ruminantium* carece de un sistema metil coenzima reductasa II (mcr II o mrt). En otros metanógenos, el cúmulo mcrII codifica una isoenzima de la enzima metil CoM reductasa I, que se regula a más durante el cultivo a altas presiones parciales de H₂ (Reeve et al., 1997). Se utiliza H₂ rápidamente en el rumen y no se acumula en altos niveles, por lo que *M. ruminantium* parece adaptarse al uso de bajos niveles de H₂ por medio del sistema mcr I únicamente.

Se identificó un total de 169 ORFs con contenido de péptidos señal en el genoma de *M. ruminantium*. (FIG. 7). De estos, se identificaron 102 péptidos señal utilizando los tres modelos SignalP y se alinearon las secuencias de aminoácidos de estos péptidos señal (FIG. 2) y se generó un logo de secuencia proteica (FIG. 3A). Se identificó una secuencia núcleo de 17 aminoácidos hidrófobos (KKIIIIIIIIIIIIIIIIISII; SEC. ID. NO: 119). SignalP-HMM calcular la probabilidad de que la secuencia contenga un péptido señal. Esta probabilidad de péptidos señal es un valor comprendido entre 0 y 1, definiéndose 0,5 como valor de corte para distinguir entre péptidos señal y péptidos no señal para este análisis. El puntaje Y de SignalP-NN da la mejor estimación del lugar donde se escinde un SP (FIG. 6). El puntaje Y se define como el promedio geométrico del puntaje C (puntaje bruto del sitio de escisión) y una pendiente allanada del puntaje S (puntaje de péptido señal) generado por SignalP-NN. El puntaje Y es un valor comprendido entre 0 y 1, donde los puntajes más altos son indicativos de una predicción satisfactoria del sitio de escisión.

Se sintetizó una secuencia de aminoácidos consenso (FIG. 3C) y se la conjugó a la marca de fluoresceína fluorescente por medio de un residuo adicional de lisina N-terminal (FIG. 3D) haciendo que la longitud final del péptido sea de 17 aminoácidos. Se analizó el péptido FITC purificado para determinar la permeabilización de las células de *M. ruminantium* (FIG. 4). En el ensayo de permeabilización celular de *M. ruminantium* 23.5% del péptido se mantenía en el sobrenadante sin unirse a las células al cabo de 30 minutos a 37°C. Se pudo retirar 3.4 % más del péptido de las células mediante 3 lavados con buffer. Se recuperó aproximadamente 62.9 % del péptido después de la extracción con SDS al 1 % de las células, indicando que la mayor parte del péptido estaba asociada a las células. Del péptido restante, se encontró que 5.8 % estaba en el interior de la fracción intracelular y 4.4 % estaba asociado a la fracción de pared/membrana celular. Por lo tanto, 4,8 % del péptido inicial (equivalente a 1.16 µg) se pudo unir a *M. ruminantium* y cruzar la membrana celular para penetrar en el citoplasma de la célula, lo que representa aproximadamente 2,3 x 10⁶ moléculas peptídicas por célula.

EJEMPLO 3: Comentarios generales

Se optó por *Methanobrevibacter ruminantium* para la secuenciación del genoma debido a su prevalencia en el rumen en una variedad de condiciones dietarias (basándose en datos de cultivo y detección molecular), la disponibilidad de los cultivos, su aptitud para el cultivo de rutina en el laboratorio, y la cantidad relativamente grande de estudios anteriores y literatura de antecedentes disponible respecto de este organismo. Se ha asignado una función a un número considerable de secuencias dentro de *M. ruminantium* y de esa manera se ha permitido obtener una imagen detallada del estilo de vida de este organismo dentro del rumen. La dependencia de *M. ruminantium* de sustratos sencillos (H₂ + CO₂, formiato) y su interacción con el medio del rumen a través de proteínas superficiales y exopolisacáridos son blancos importantes para la inhibición. Los datos de secuencias dilucidan el metabolismo de este organismo y la manera en que interactúa con otros microbios y apunta a los sistemas conservados y componentes entre los metanógenos, que se pueden inactivar para prevenir o reducir la formación de metano en el rumen.

45 Referencias

- Altermann E, Klaenhammer TR (2005) Trayectoria Voyager: mapeo de trayectoria utilizando la base de datos de Enciclopedia de Genes y Genomas de Kyoto (KEGG) (Pathway Voyager: pathway mapping using the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. *BMC Genomics* 6:60-66.
- Altermann, E., y T. R. Klaenhammer. 2003. GAMOLA: una nueva solución local para anotación de secuencias y análisis de genomas procariotas preliminares y terminados (GAMOLA: a new local solution for sequence annotation and analyzing draft and finished prokaryotic genomes). *Omics* 7:161-169.
- Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) BLAST con Brecha y PSI-BLAST: nueva generación de programa de búsqueda de base de datos de proteína (Gapped BLAST and PSI-BLAST: a new generation of protein database search programs). *NucleicAcids Research* 25, 3389-3402.
- Balch WE, Fox GE, Magrum LJ, Woese CR, Wolfe RS (1979) Metanógenos: reevaluación de un grupo biológico único (Methanogens: reevaluation of a unique biological group). *Microbiological Reviews* 43, 260-296.
- Baresi, L. y Bertani, G. 1984. Aislamiento de un bacteriófago para una bacteria metanogénica. (Isolation of a bacteriophage for a methanogenic bacterium). En Abstracts of the Annual Meeting of the American Society for Microbiology. Washington DC: American Society for Microbiology, pag. 133.
- Bendtsen, J.D., Nielsen, H., von Heijne, G. y Brunak, S. 2004. Anticipación mejorada de péptidos de señal: SignalP 3.0(Improved prediction of signal peptides) *J Mol Biol.* 16(4):783-95.

- Bickle, T. A. y D. H. Kruger. 1993. Biología de restricción de AND (Biology of DNA restriction). *Microbiol. Rev.* 57:434-450.
- Bult CJ, y otros (1996). Secuencia de genoma complete del arqueón metanogénico, *Methanococcus jannaschii*. (Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*). *Science* 273, 1058-1073.
- 5 Coutinho PM, Henrissat B (1999). Enzimas activas-carbohidrato: un método de base de dato integrado (Carbohydrate-active enzymes: an integrated datábase approach). In 'Recent Advances in Carbohydrate Bioengineering' (Eds HJ Gilbert, G Davies, B Henrissat y B Svensson) pags. 3-12 (The Royal Society of Chemistry, Cambridge) (Carbohydrate Active Enzymes database, protocolo de transferencia de hipertexto://world wide web.cazy.org/).
- 10 Delcher AL, Harmon D, Kasif S, White O, Salzberg SL (1999). Identificación mejorada de gen microbiano con GLIMMER (Improved microbial gene identification with GLIMMER). *Nucleic Acids Research* 27, 4636-4641.
- Emanuelsson, O., Brunak, s., von Heijne, G. y Nielsen, H. 2007. Localización de proteínas en la célula utilizando Targetp Signalp y herramientas relacionadas (Locating proteis i te cell usin TargetP, SignalP and related tolos). *Nat Protoc.* 2(4):953-71.
- 15 Fleischmann y otros, 1995. Secuenciación aleatoria de genoma complete y ensamble de *Haemophilus influenza* (Whole-genome random sequencing and assembly of *Haemophilus influenza*) *Rd Science* 269:496-512.
- Fricke WF, Seedorf H, Henne A, Kruer M, Liesegang H, Hedderich R, Gottschalk G, Thauer RK (2006). La secuencia de genoma de *Methanosphaera stadtmanae* revela porqué este arqueón intestinal humano es restringido a metabol y H₂ para la formación de metano y síntesis de ATP. (The genome sequence of *Methanosphaera stadtmanae* reveals why this human intestinal archaeon is restricted to methanol and H₂ for methane formation and ATP synthesis). *Journal of Bacteriology* 188, 642-658.
- 20 Galagan et al. 2002. El genoma de *Methanosarcina acetivorans* revela diversidad metabólica y fisiológica extensa (The genoma of *Methanosarcina acetivorans* reveals extensive metabolic and physiological diversity). *Genome Res.* 12:532-542.
- 25 Godde JS, Bickerton A (2006). El ADNe repetitivo denominado CRISPRs y sus genes asociados: evidencia de transferencia horizontal entre procariotos (The repetitive DNAe called CRISPRs and their associated genes: evidence of horizontal transfer among prokaryotes). *Journal of Molecular Evolution* 62, 718-729.
- Haft DH, Selengut J, Mongodin EF, Nelson KE (2005). Una comunidad de 45 familias de proteína asociadas con CRISPR (Cas) y múltiples sub-tipos CRISPR/Cas existen en genomas procariotos (A guild of 45 CRISPR-associated
- 30 (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes). *PLoS Computational Biology* 1:474-483.
- Hamilton, P. T. y J. N. Reeve. 1985. Divergencia de secuencia de un gen arqueobacteriano clonado de un metanogen mesofílico y uno termofílico (Sequence divergence of an archaeobacterial gene cloned from a mesophilic and a thermophilic methanogen) *J. Mol. Evol.* 22:351-360.
- 35 Jansen R, Embden JD, Gaastra W, Schouls LM (2002). Identificación de genes que están asociados con repeticiones de AND en procariotos (Identification of genes that are associated with DNA repeats in prokaryotes). *Molecular Microbiology* 43, 1565-1575.
- Jansen R, van Embden JD, Gaastra W, Schouls LM (2002). Identificación de una familia novedosa de repeticiones de secuencia entre procariotos (Identification of a novel family of sequence repeats among prokaryotes). *OMICS: A*
- 40 *Journal of integrative biology* 6, 23-33.
- Jensen, L. J., Früs, C. y Ussery, D. W. 1999. Tres vistas de genomas microbianos (Three views of microbial genomes). *Res. Microbiol.* 150, 773-777.
- Joblin KN, Naylor GE, Williams AG (1990). Efecto de *Methanobrevibacter smithii* en la actividad xilanolítica de hongo de rumiante anaeróbico (Effect of *Methanobrevibacter smithii* on xylanolytic activity of anaerobic ruminal fungi. *Applied and Environmental Microbiology*) 56, 2287-2295.
- 45 Kanehisa M, Goto S, Kawashima S, Okuno Y, Hattori M (2004). El recurso KEGG para descifrar el genoma (The KEGG resource for deciphering the genome). *Nucleic Acids Research* 32, D277-D280.
- Kiener, A., Konig, H., Winter, J. y Leisinger, T. 1987. Purificación y uso de endopeptidasa de pseudomureína de *Methanobacterium wolfei* para lisis de *Methanobacterium thermoautotrophicum*. (Purificación and use of
- 50 *Methanobacterium wolfei* pseudomurein endopeptidase for lysis of *Methanobacterium thermoautotrophicum*). *J. Bacteriol.* 169, 1010-1016.
- Knox, M. R. y Harris, J. E. 1986. Aislamiento y caracterización de un bacteriofágo de *Methanobrevibacter smithii*. (Isolation and characterisation of a bacteriophage of *Methanobrevibacter smithii*). En Abstracts of the XIV International Congress on Microbiology. Manchester: International Union of Microbiological Societies.
- 55 Kurtz S, Schleiermacher C (1999) REPuter: cómputo rápido de repeticiones máximas en genomas completos (REPuter: fast computation of maximal repeats in complete genomes). *Bioinformatics* 15, 426-427.
- Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin F., Wallace, i.m., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J. y Higgins, d.G. Clustal W y clustal X versión 2.0. *bioinformatics* 23(21):2947-8.
- 60 Lowe TM, Eddy SR (1997) tRNAscan-SE: programa para la detección mejorada de genes de transferencia de ARN en secuencia genómica (tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence). *Nucleic Acids Research* 25, 955-964.
- Loenen, W. y N. Murray. 1986. Mejoría de modificación mediante la proteína de alivio de restricción (Ra1)

- bacteriofago lambda (Modification enhancement by restriction alleviation protein (Ral) of bacteriophage lambda). *J. Mol. Biol.* 190:11-22.
- Lucchini, S., F. Deslere, y H. Brussow. 1999. Genómicos comparativos de especies de fago *Streptococcus thermophilus* soporta una teoría de evolución modular (Comparative genomics of *Streptococcus thermophilus* phage species supports a modular evolution theory). *J. Viral.* 73:8647-8656.
- 5 Luo, Y. N., Pfister, P., Leisinger, T. y Wasserfallen, A. 2002. Endoisopeptidasas de pseudomureína PeiW y PeiP, dos miembros moderadamente relacionados de una familia novedosa de proteasas producidas en cepas *Methanothermobacter* (Pseudomurein endoisopeptidasas PeiW and PeiP, two moderately related members of a novel family of proteases produced in *Methanothermobacter* strains). *FEMS Microbiol. Lett.* 208, 47-51.
- 10 Makarova, K. S., Aravind, L. y Koonin, E. V. 1999. Superfamilia de proteínas asqueales, bacterianas y eucariotas homólogas a transglutaminasas animales (A superfamily of archaeal, bacterial, and eukaryotic proteins homologous to animal transglutaminases). *Protein Sci.* 8, 1714-1719.
- Makarova KS, Grishin NV, Shabaiina SA, Wolf YI, Koonin EV (2006). Sistema inmune a base de interferencia-ARN putativo en procariontes: Análisis de cómputo de la maquinaria enzimática anticipada, analogías funcionales con ARNi eucarioto, y mecanismos de acción hipotéticos (A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action). *Biology Direct* 1:7-32.
- 15 Nueva Zelanda Statistics 2005 (www.stats.govt.nz)
- New Zealand's Greenhouse Gas Inventory 1990 - 2004. Reporte de Inventario Nacional y Formato de Reporte Común (The National Inventory Report and Common Reporting Format). (2006) Ministry for the Environment. Protocolo de transferencia de hipertexto://www.mfe.govt.nz/publications/climate/nir-apr06/nir-apr06.pdf.
- 20 Nielsen, H. Engelbrecht, J. Brunak S. y von Heijne, G. 1997. Identificación de péptidos de señal procariontes y eucariotes y anticipación de sus sitios de disociación (Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites) *Prot. Eng.*, 10:1-6
- 25 Ono, M., Wada, Y., Wu, Y., Nemori, R., Jinbo, Y., Wang, H., Lo, K. M., Yamaguchi, N., Brunkhorst, B., Otomo, H. et al. (1997) FP-21399 bloquea la fusión de membrana transmitida por proteína de envoltura HIV y concentrados en nodos linfáticos (FP-21399 blocks HIV envelope protein mediated membrane fusion and concentrates in Lymph nodes.) *Nat. Biotechnol.* 15, 343-348.
- Pérez-Bercoff, a., Koch, J. y Bürglin, t. T. 2006. Logobar: visualización de gráfica de barras de logos de proteína con brechas (LogoBar: bar graph visualization of protein logos with gaps) *Bioinformatics* 22(1):112-4
- 30 Rawlings, N. D., Morton, F. R. y Barrett, A. J. 2006. MEROPS: la base de datos de péptidasa (MEROPS: the peptidase database). *Nucleic Acids Res.* 34, D270-D272.
- Reeve JN, Nolling J Morgan RM, Smith DR (1997) Metanogénesis: genes, genomas y quién es el primero? (Methanogenesis: genes, genomes and who's on first?) *Journal of Bacteriology* 179, 5975-5986.
- 35 Samuel BS, Hansen EE, Manchester JK, Coutinho PM, Henriessat B, Fulton R, Latreille P, Kim K, Wilson RK, Gordon JI (2007). Adaptaciones genómicas de *Methanobrevibacter smithii* para el intestino humano (Genomic adaptations of *Methanobrevibacter smithii* to the human gut). *Proceedings of the National Academy of Sciences USA* 104, 10643-10648.
- Salzberg et. al., 1998. Identificación de gen microbiano utilizando modelos de Markov interpolados (Microbial gene identification using interpolated Markov models) *Nucleic Acids res.* 26:544-8.
- 40 Smith et. al., (1997). Secuencia de genoma completo de *Methanobacterium thermoautotrophicum* ΔH: Análisis funcional y genómicos comparativos (Complete genome sequence of *Methanobacterium thermoautotrophicum* ΔH: Functional analysis and comparative genomics. *Journal of Bacteriology* 179, 7135-7155.
- Smith PH, Hungate RE (1958). Aislamiento y caracterización de especie y *Methanobacterium ruminantium* n. sp. (Isolation and characterization of *Methanobacterium ruminantium* n. sp.) *Journal of Bacteriology* 75, 713-718.
- 45 Staden R, Beal KF, Bonfield JK (1998). El Paquete Staden. Métodos en Biología Molecular: Métodos y Protocolos de Bioinformática (The Staden Package. *Methods in Molecular Biology: Bioinformatics Methods and Protocols*) 132, 115-130.
- Tatusov RL, Natale DA, Garkavtsev IV, Tatusova TA, Shankavaram UT, Rao BS, Kiryutin B, Galperin MY, Fedorova ND, Koonin EV (2001). La base de datos COG: nuevos desarrollos en clasificación filogenética de proteínas de genomas completos (The COG database: new developments in phylogenetic classification of proteins from complete genomes). *Nucleic Acids Research* 29, 22-28.
- 50 von Heijne, G. 1985. Secuencias de señal: límites de variación (Signal sequences: The limits of variation *J. Mol. Biol.* 184-99-105
- 55 En los casos en que, en la descripción precedente, se haya hecho referencia a enteros con equivalentes conocidos de los mismos, los equivalentes se incorporan a la presente como si se los hubiera expuesto en forma individual. Si bien se ha descrito la invención en relación con modalidades preferidas específicas, se ha de entender que la invención no debe limitarse indebidamente a dichas modalidades específicas.

REIVINDICACIONES

1. Un polipéptido o péptido aislado: a) que comprende SEC. ID. NO: 117, 118 o 119; b) que es una variante funcional de SEC. ID. NO: 117, y comprende una secuencia de aminoácidos que comparte por lo menos el 90% de la identidad con la secuencia entera de SEC. ID. NO: 117; c) que es una variante funcional de SEC. ID. NO: 118, y comprende una secuencia de aminoácidos que comparte por lo menos el 95% de la identidad con la secuencia entera de SEC. ID. NO: 118; d) que es una variante funcional de SEC. ID. NO: 119, y comprende una secuencia de aminoácidos que comparte por lo menos el 90% de la identidad con la secuencia entera de SEC. ID. NO: 119; e) un fragmento de SEC. ID. NO: 117, cuyo fragmento incluya por lo menos 15 aminoácidos de SEC. ID. NO: 117; f) un fragmento de SEC. ID. NO: 118, cuyo fragmento incluya por lo menos 15 aminoácidos de SEC. ID. NO: 118; g) un fragmento de SEC. ID. NO: 119, cuyo fragmento incluya por lo menos 15 aminoácidos de SEC. ID. NO: 119, en el que dicho polipéptido o péptido tiene actividad de asociación celular o permeabilización celular.
2. Un polipéptido aislado que comprende: a) una secuencia codificante entera de SEC. ID. NO: 117, 118 o 119; b) una secuencia codificante para el péptido o polipéptido que comparte por lo menos el 90% de identidad con la secuencia entera de SEC. ID. NO: 117; c) una secuencia codificante para el péptido o polipéptido que comparte por lo menos el 95% de identidad con la secuencia entera de SEC. ID. NO: 118; d) una secuencia codificante para el péptido o polipéptido que comparte por lo menos el 90% de identidad con la secuencia entera de SEC. ID. NO: 119; e) una secuencia codificante del fragmento de SEC. ID. NO: 117, 118 o 119, cuyo fragmento incluya por lo menos 15 aminoácidos de SEC. ID. NO: 117, 118 o 119; f) SEC. ID. NO: 511, 512 o 513; o g) una secuencia nucleotídica complementaria a cualquiera de las (a) a (f), en la que dicho polinucleótido codifica para un péptido o polipéptido con actividad de asociación celular o permeabilización celular.
3. Un vector que comprende el polinucleótido aislado de la reivindicación 2.
4. Una célula huésped a) que está modificada genéticamente para codificar el polipéptido o péptido de la reivindicación 1; b) que está modificada para comprender el polinucleótido de la reivindicación 2; c) que comprenda el vector de la reivindicación 3.
5. La célula huésped de la reivindicación 4, que es seleccionada de una célula procariota y una célula metanógena.
6. La célula huésped de la reivindicación 5, que se selecciona de *Escherichia coli* y *Methanobrevibacter ruminantium*, incluyendo la cepa M1T (DSM1093) *Methanobrevibacter ruminantium*.
7. Una molécula conjugada o de fusión que comprende a) el polipéptido o péptido de la reivindicación 1; o b) el polinucleótido de la reivindicación 2.
8. Una molécula conjugada o de fusión que comprende el polipéptido o péptido de la reivindicación 1, y que además comprende un compuesto anti-metanogénico, un anticuerpo o un fragmento de anticuerpo, un enzima lítico, un ácido nucléico de un péptido, un péptido antimicrobiano, o un antibiótico.
9. Un método para la permeabilización de una célula microbiana que comprende: contactar con la célula con el polipéptido o péptido de la reivindicación 1.
10. Un método para la permeabilización de una célula microbiana que comprende: contactar con la célula con la molécula conjugada o de fusión de la reivindicación 7 o de la reivindicación 8.
11. El método de la reivindicación 9 o 10, en el cual la célula es metanógena.
12. El método de la reivindicación 11, en la que la célula es *Methanobrevibacter ruminantium*.
13. El método de la reivindicación 12, en la que la célula es de la cepa M1T (DSM1093) de *Methanobrevibacter ruminantium*.

FIG. 1A

Comparación de genomas Metanobacterianos

Metanógeno	Mb	ORFs	% G+C	ARNr	ARNt
<i>Methanobrevibacter ruminantium</i> M1 ^a	2,9	2239	32,6	2	59
<i>Methanobrevibacter smithii</i> PS ^b	1,9	1795	31,0	2	34
<i>Methanobrevibacter thermoautotrophicus</i> ΔH ^c	1,8	1873	49,5	2	39
<i>Methanobrevibacter</i> stadtmannae DSM3091 ^d	1,8	1534	27,6	4	40

^a El tamaño del genoma y el número de ORFs se basan en el análisis de la secuencia preliminar del genoma de *M. ruminantium* de un solo contiguo

^b Samuel et al., 2007

^c Smith et al., 1997

^d Fricke et al., 2006

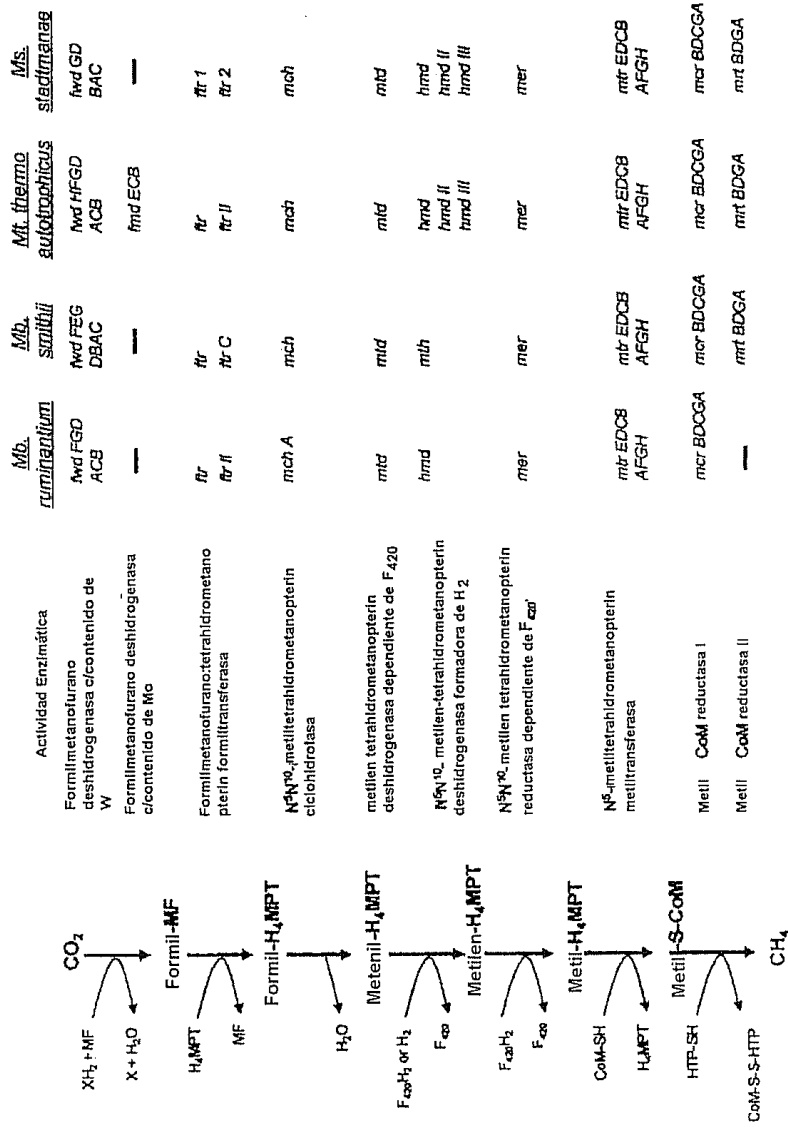
FIG. 1B

Estadísticas del genoma preliminar de *M. ruminantium*

Tamaño del genoma (pb)	2937347
Marcos de lectura abiertos	2239
Proteínas con dominios transmembrana	503 (22,5)
Estructuras de terminación	334 (14,9)
TIGRfams	2304
Pfams	3315
COGs	1834

^a Los números entre paréntesis indican la característica en términos de % del número total de ORF

FIG. 1C



Actividad Enzimática	<i>Mb. ruminantium</i>	<i>Mb. smithii</i>	<i>Mt. thermophilus</i>	<i>Ms. sibirianae</i>
Formilmetanolurano deshidrogenasa c/contenido de W	<i>fmf FGD ACB</i>	<i>fmf FEG DBAC</i>	<i>fmf HFCD ACB</i>	<i>fmf GD BAC</i>
Formilmetanolurano deshidrogenasa c/contenido de Me	—	—	<i>fmf ECB</i>	—
Formilmetanolurano:tetrahidrometano pterin formiltransferasa	<i>fr fr II</i>	<i>fr fr C</i>	<i>fr fr II</i>	<i>fr fr 1 fr 2</i>
N ⁵ N ¹⁰ -metiltetrahidrometanopterin ciclohidrolasa	<i>mch A</i>	<i>mch</i>	<i>mch</i>	<i>mch</i>
metilten tetrahidrometanopterin deshidrogenasa dependiente de F ₄₂₀	<i>mid</i>	<i>mid</i>	<i>mid</i>	<i>mid</i>
N ⁵ N ¹⁰ -metilten-tetrahidrometanopterin deshidrogenasa formadora de H ₂	<i>himd</i>	<i>him</i>	<i>himd II himd III</i>	<i>himd II himd III</i>
N ⁵ N ¹⁰ -metilten tetrahidrometanopterin reductasa dependiente de F ₄₃₀	<i>mer</i>	<i>mer</i>	<i>mer</i>	<i>mer</i>
N ⁵ -metiltetrahidrometanopterin metiltransferasa	<i>mtr EDCB AFGH</i>	<i>mtr EDCB AFGH</i>	<i>mtr EDCB AFGH</i>	<i>mtr EDCB AFGH</i>
Metil CoM reductasa I	<i>mcr BDCGA</i>	<i>mcr BDCGA</i>	<i>mcr BDCGA</i>	<i>mcr BDCGA</i>
Metil CoM reductasa II	—	<i>mtt BDGA</i>	<i>mtt BDGA</i>	<i>mtt BDGA</i>

FIG. 2-1

NOMBRE DEL GEN	SECuencia SENAL	SEC. ID NO:
Contig40_gen_70	-----MKEKIIISILVIAIAISVITPAFS-----	9
Contig49_gen_208	-----MDKKIIGAVVALLVI--IVGA-----	162
Contig40_gen_867	-----MREKILIAAIAIILIL--CGGVFA-----	65
Contig47_gen_125	-----MDKKMIVSVAFLILLI--AVALVSVFDEENS-----	130
Contig47_gen_391	-----MDKKMTVLLVALRCHL--CVG--S-----	146
Contig40_gen_34	-----MVALASIIILISSIAAASA-----	1
Contig45_gen_106	-----MKEKNSHILIVSLISIFLLISIAASA-----	121
Contig45_gen_36	-----MFKVSKLIVCVLSVSLFLLVQASA-----	111
Contig55_gen_23	-----MLLNDKSEILKSLISILFLLVITLITFNSVIA-----	170
Contig40_gen_231	-----MKKNLSKNILLISLIFLFLVLSIGSFA-----	23
Contig40_gen_1099	-----MNEFKLLMI SLLLFVLSVGFSTASA-----	87
Contig40_gen_1254	-----MKENSRVIGILSLFLFVITLIVSVGA-----	96
Contig40_gen_39	-----MDN--KKIFVIVALAI--LAIIVAGSVS--A-----	3
Contig40_gen_40	-----MNN--KKIFAIAAALAI--LAIIVAGSVS--A-----	4
Contig40_gen_41	-----MNA--KLLTILAAALAI--LAIIVAGSVS--A-----	5
Contig40_gen_1036	-----MNN--KKIFV--AGLAI--LAIIVAGSVS--A-----	79
Contig40_gen_732	-----MNS--KKLIVTALAI--LAIIVAGSVS--A-----	47
Contig40_gen_1215	-----MDS--KKIIMIAVVAL--LAIIVAGSVS--A-----	93
Contig47_gen_146	-----MDS--KKLIVTALAI--LAIIVAGSVS--A-----	132
Contig40_gen_964	-----MSI--KRILLSLMLF--IIFISIFVS--A-----	71
Contig40_gen_1274	-----MKN--KSLIISLILL--IIFISIFVS--A-----	98
Contig40_gen_1350	-----MN--KKIISLHVL--IVALSIVSAYA--A-----	101
Contig40_gen_1355	-----MNN--KKIIMSFLVL--LALSIVSAYS--A-----	103
Contig47_gen_381	-----MRF--KKYLFILLIAL--ICIIISVSAYA-----	145
Contig40_gen_203	-----MKNNLKKTTIILALMALLIISIGALS--A-----	21
Contig40_gen_837	-----MK--LRKFSVILAVLL--VAIIAIGAYS--A-----	61
Contig40_gen_553	-----MK--KKIAILGIAL--LAFVIGASS--A-----	43
Contig45_gen_116	-----MNS--KKIIVLGIIL--LSEFAIVGSAS--A-----	122
Contig40_gen_834	-----MNS--NKTYAVLGLLL--LIIISIGALS--A-----	58
Contig40_gen_105	-----MNNLKKTFELGIVLIGLISFNLS--A-----	12
Contig55_gen_45	-----MKINLRV--ILGIIILIIICISSAIIIS--A-----	172
Contig40_gen_228	-----MNSKGYL--VLFLLIILIS--FSIISASF--A-----	22

FIG. 2-2

Contig40_gen_269	-----M--KRRYK-VLFLLAALLT-IISINPAIS-A-----	29
Contig40_gen_815	-----MILAAILL-AVGMTLGAVS-A-----	55
Contig49_gen_184	-----MKNKAMFLISALLI-AVILLSLGAVS-AAADDAIA-----	160
Contig40_gen_63	-----MKVYQLSSILAIVL-ILFLSLAVS-A-----	8
Contig40_gen_187	-----MFKRKAVALAISILA-VIFASMCIVS-A-----	20
Contig47_gen_197	-----MLISVLG-VIVIIIMVVA-A-----	134
Contig40_gen_352	-----MKK-SVFKILITALA-LILLAVSIVS-S-----	36
Contig47_gen_13	-----MNKQVVF-ALILLT-IILLSVVAVS-----	125
Contig40_gen_841	-----MILLSLI-LVILSISCVS-A-----	62
Contig49_gen_239	-----MLIMAI-LVILLTMAVS-ASELEDIQVT-----	164
Contig40_gen_1074	-----MEKTRKSKLFIILLII-SLISISVS-A-----	80
Contig40_gen_1238	-----MK--LKSRYFVLLII-CILFISITVS-A-----	94
Contig47_gen_7	-----MKRNFVGVFIV-FILFISITVS-A-----	124
Contig40_gen_232	-----MKRNIYFIILLVT-LFLISMVVS-A-----	24
Contig40_gen_930	-----MKPLIIKISIFCNKRIFFITMI-VMLLSLAAS-A-----	70
Contig49_gen_175	-----MLNKALISLIV-LFMLSISAVS-A-----	157
Contig40_gen_1084	-----MD-KKIFIVSFIL-LAIFTIGAVG-A-----	81
Contig40_gen_1089	-----MVTMNRKLFIVSLII-ILILITIGAVS-A-----	83
Contig40_gen_1088	-----MMKMKRKNLELISLIL-LIILITIGAVS-A-----	82
Contig49_gen_167	-----MNNKTLFIFLFI-CILFTFIMVS-A-----	155
Contig40_gen_824	-----MKRSL-----FLYIALIFT-ISLISFSAVS-A-----	56
Contig47_gen_375	-----LTFNNLRINIDCMVIRVVF-VLLLSILAVS-A-----	144
Contig40_gen_1158	-----MKVLK-----IAIIML-ILIIISLGAVS-A-----	90
Contig49_gen_43	-----MRLRY-----FAIISI-ILIIFLVPSFA-----	150
Contig45_gen_91	-----MEIRYRNLKVTIFL-VMLISCGFAS-A-----	115
Contig45_gen_93	-----MKIRYRNLKVTIFL-VMLISCGFAS-A-----	116
Contig40_gen_164	-----MLIVTTCIVLIIIVLYFLFPEGLTNS-----	16
Contig40_gen_1364	-----MKIQRGYIILITLVVLF--LSAASAA-----	104
Contig47_gen_57	-----MLNKKIILITFILLS--ISSASASA-----	126
Contig55_gen_40	-----MKKILGTCTILFIL--ISVAYA-----	171
Contig40_gen_1029	-----MRNPKDYMKVDYLIIMMALLIS--IVSPIAA-----	78
Contig40_gen_169	-----LKSDKRAKFAIEFSIAILAIG--LSNIAA-----	18
Contig47_gen_304	-----LS-----VILLFLA--VSTVAA-----	138
Contig40_gen_346	-----MEDNLLKRNKLIISIFLVSILA--ISAVSA-----	34
Contig40_gen_872	-----MLISIVLISHIA--IGAVSA-----	66
Contig40_gen_750	-----MISLHLSHIA--ISAASAA-----	50
Contig40_gen_1008	-----MILISLFLSILA--IGAASA-----	75
Contig40_gen_501	-----MLNKKFIISIIIIIFIS--ISAISA-----	42

FIG. 2-3

Contig40_gen_1021	-----MKLYKNSIIILILILILS--IGRAAA-----	76
Contig40_gen_721	-----MK-----RSIIFFTLISLFV--IGYASA-----	45
Contig40_gen_1093	-----MKFNKRGISAIISIIILILS--ISMASA-----	84
Contig40_gen_1097	-----VFILKFEIKRSLIFSIILILS--IGMASA-----	85
Contig40_gen_1176	-----MNFKTKGSLIISLILFIIIG--IGMASA-----	91
Contig40_gen_72	-----MNRKFKLLIFIFIAFALINTCFIINDNLS-----	10
Contig49_gen_25	-----MRKILFLTIMILICTINSVCA-----	149
Contig40_gen_411	-----M-RKNIFLIAIILIAVAVSGC-----	37
Contig49_gen_152	-----MDKYLAIATVIAIVAVG-----	154
Contig40_gen_636	-----MNFKIIILIAVFIASVGI VAA-----	44
Contig40_gen_1351	-----MSLSIFVLVIGGFINKRIILIFVLIFFISGSVVA-----	102
Contig49_gen_194	-----MKF-----	161
Contig40_gen_909	-----MKWKLIGLIIILIAVSV--SG-----	69
Contig40_gen_996	-----MKISRII-LIILLVVFFEIGLFSSTIVNA-----	74
Contig40_gen_835	-----MINKRIISLIIILVFLIIG-LSAVSA-----	59
Contig40_gen_349	-----MNRN-----KIIVLVLLIAVGFIMGPCA-----	35
Contig47_gen_140	-----MKIS--RIIVLMIIL-----FTAGMVA-----	131
Contig40_gen_1104	-----MKIKKS--FVILCLIIICL-----FTIASVAA-----	88
Contig47_gen_62	-----VFSVSLNKLKIG--RVFCLFIIH--FISCSINCVFA-----	128
Contig40_gen_828	-----MKYN--KKIIFLIIICL--IIPQAIY-----	57
Contig40_gen_1247	-----MNYN--IITFIIFLMDAL--VLMASIQVCGA-----	95
Contig40_gen_1198	-----MG--KFKFIIIVLAL--FLICGIA-----	92
Contig40_gen_1098	MDALIPVKDNELLVTKKSDFK--RIFICLVLLTCL--IGAVSA-----	86
Contig49_gen_44	-----MFIGLIIIGLL--IIPISFA-----	151
Contig40_gen_836	-----MD-FKK--AIPLFALLIIL--FIIGSSA-----	60
Contig49_gen_183	-----MSYFNKGHITWIIICLLIGTL--AMMGSASA-----	159
Contig40_gen_847	-----MDNSNI--ISVIIVLCIAAGVTA-----	63
Contig40_gen_848	-----MDNSSII--ISVIIVLCIAAGVTA-----	64
Contig40_gen_766	-----MKTKICGIIKRNPEIMLAAGVLSHA-----	52
Contig45_gen_64	-----MKIIVAGVGVGLSLAVLLAOKHDVTA-----	113
Contig45_gen_89	-----MNMKIIIVAGVGVGLSIIALLAOKHDVTA-----	114
Contig55_gen_4	-----MFLVILLFAFIVIGGSYSVEA-----	167

FIG. 3A

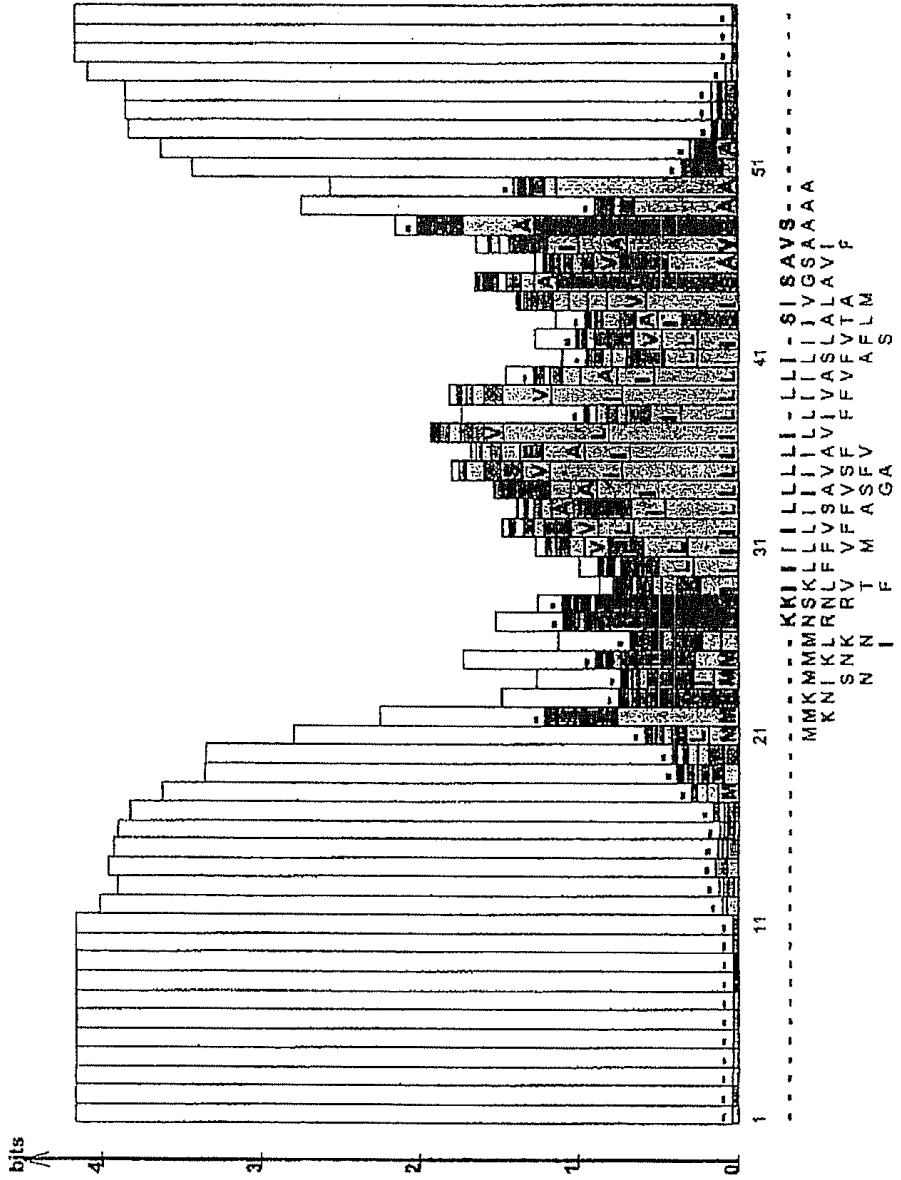


FIG. 3B

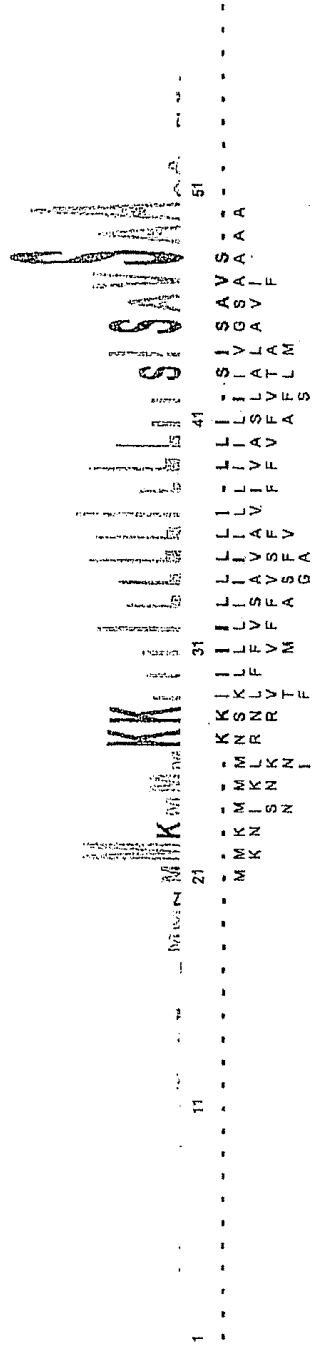


FIG. 3C

KKLIILLIILLIILLI
SEC ID NO:117

FIG. 3D

Fluoresceína -KKLIILLIILLIILLI
SEC ID NO:118

FIG. 4

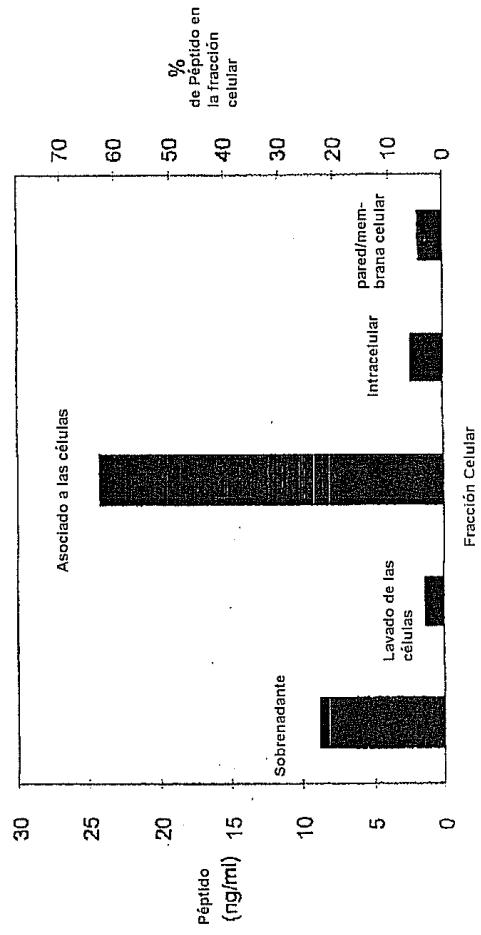


FIG. 5

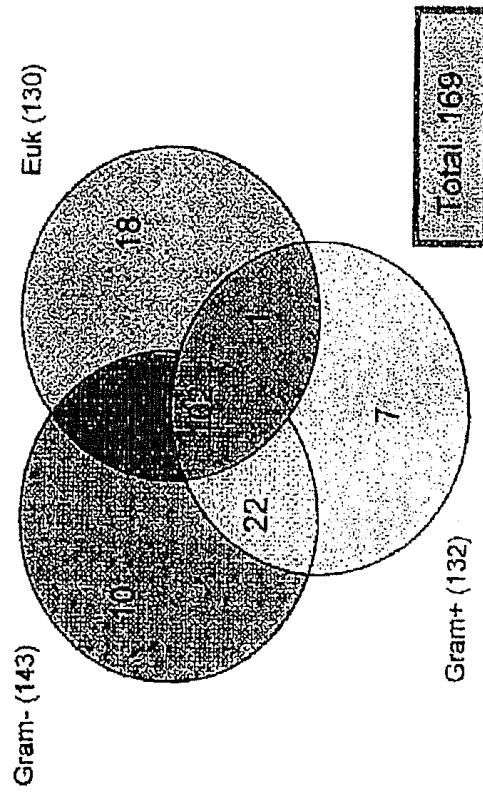


FIG. 6-1

Número del gen Sitio de inserción	Anotación	Gram+	Gram-	Euk	Puntaje de probabilidad de SignalP-HMM SP	Gram+	Gram-	Euk	Puntaje Y & previsto SignalP-NN
Contig40_gene_34	Proteína hipotética	0.931	0.998	1	0.635	19	0.534	19	0.886
Contig40_gene_35	Proteína de la familia LemA	0.839	0.635	0.991	0.122	36	0.083	15	0.472
Contig40_gene_39	Proteína hipotética	0.999	1	0.991	0.736	26	0.777	26	0.748
Contig40_gene_40	Proteína hipotética	1	1	0.999	0.698	26	0.704	26	0.819
Contig40_gene_41	Proteína hipotética	1	1	0.999	0.601	26	0.745	26	0.8
Contig40_gene_51	Proteína tipo adhesina	0.017	0.016	0.938	0.433	26	0.338	26	0.448
Contig40_gene_54	Proteína hipotética	0.848	0	0.957	0.363	25	0.162	19	0.631
Contig40_gene_56	Proteína tipo adhesina	1	1	0.997	0.775	26	0.621	26	0.92
Contig40_gene_70	Proteína hipotética	1	0.999	0.882	0.807	25	0.633	25	0.789
Contig40_gene_72	Proteína hipotética	1	0.971	0.59	0.612	32	0.406	18	0.484
Contig40_gene_75	Proteína hipotética	0.048	0.143	0.996	0.338	35	0.2	31	0.41
Contig40_gene_105	Proteína tipo adhesina	1	1	0.816	0.789	29	0.739	29	0.726
Contig40_gene_119	proteína A de biosíntesis de melibiodotriin guanina dinucleótido	MobA	0.007	0.582	0.483	24	0.279	26	0.39
Contig40_gene_161	Proteína hipotética	0.914	0.982	0.174	0.278	34	0.199	34	0.439
Contig40_gene_163	2-deshidropantoato 2-reductasa PAN E	0.736	0.941	0.114	0.323	25	0.161	14	0.249
Contig40_gene_164	Proteína hipotética	0.824	0.828	0.675	0.362	28	0.469	28	0.843
Contig40_gene_165	Proteína hipotética	0.292	0.36	0.382	0.348	26	0.388	26	0.624
Contig40_gene_169	Proteína hipotética	0.991	0.998	0.958	0.448	28	0.314	20	0.462
Contig40_gene_179	Proteína hipotética	0.709	0.922	0.128	0.38	44	0.119	44	0.149
Contig40_gene_187	Proteína tipo adhesina	1	1	0.994	0.766	26	0.551	26	0.699
Contig40_gene_203	Proteína tipo adhesina	1	1	0.971	0.652	29	0.684	29	0.848
Contig40_gene_228	Proteína de contenido dominio SlnA	1	1	0.93	0.717	27	0.734	27	0.696
Contig40_gene_231	Proteína tipo adhesina	1	1	0.98	0.677	29	0.814	29	0.699
Contig40_gene_232	Proteína tipo adhesina con dominio estirina proteasa	1	1	0.989	0.777	25	0.704	25	0.802
Contig40_gene_248	Proteína hipotética	0.721	0.346	0.112	0.244	22	0.295	22	0.185
Contig40_gene_251	Proteína hipotética	0.769	0.931	0.099	0.514	26	0.635	21	0.939
Contig40_gene_252	Proteína hipotética	0.408	0.929	0.991	0.603	32	0.851	32	0.85
Contig40_gene_260	Proteína hipotética	0.014	0.457	0.854	0.628	28	0.428	28	0.384
Contig40_gene_269	Proteína tipo adhesina	1	1	0.988	0.48	24	0.546	25	0.826
Contig40_gene_296	Proteína hipotética	0.605	0.045	0.381	0.207	25	0.085	25	0.407
Contig40_gene_297	Proteína hipotética	0.999	0.99	0.372	0.385	26	0.827	15	0.491
Contig40_gene_306	UDP-glucosa 4-epimerasa GAIU	0.022	0.107	0.661	0.494	36	0.458	28	0.427
Contig40_gene_317	Proteína de familia geranyl reductasa	0.458	0.954	0.446	0.808	27	0.451	27	0.746
Contig40_gene_346	Proteína tipo adhesina	1	1	0.834	0.499	30	0.755	30	0.851
Contig40_gene_349	Proteína hipotética	0.999	1	0.994	0.939	27	0.947	27	0.647
Contig40_gene_352	Proteína hipotética	1	1	0.995	0.5	25	0.402	25	0.843
Contig40_gene_411	Proteína hipotética	0.881	0.986	0.973	0.406	32	0.184	20	0.748
Contig40_gene_431	Peptidasa scial I	0.426	0.855	0.999	0.466	25	0.302	25	0.744
Contig40_gene_448	familia de peptidasa S49	0.875	0.884	0.003	0.26	30	0.086	22	0.61
Contig40_gene_466	Proteína hipotética	0.966	0.488	0.044	0.715	32	0.321	32	0.549
Contig40_gene_483	Proteína ABC transportadora de unión al sustrato	1	1	0.999	0.346	17	0.471	16	0.862
Contig40_gene_501	Proteína tipo adhesina	1	1	0.97	0.632	25	0.716	24	0.921
Contig40_gene_553	Proteína ABC transportadora de unión al sustrato	1	1	0.993	0.671	25	0.564	25	0.59

FIG. 6-2

Contig40_gene_636	proteína hipotética	1	0.981	0.613	0.749	24	0.926	25
Contig40_gene_721	Proteína ABC transportadora de unión al sustrato	0.991	0.988	0.756	24	0.851	0.907	24
Contig40_gene_730	Proteína c/contenido dominio CBC	0.026	0.862	0.232	24	0.248	0.248	33
Contig40_gene_732	proteína hipotética	0.999	0.996	0.754	26	0.68	0.771	26
Contig40_gene_733	proteína hipotética	0.448	0.881	0.455	36	0.396	0.432	25
Contig40_gene_749	proteína hipotética	0.251	0.988	0.284	31	0.281	0.353	21
Contig40_gene_750	proteína tipo adhesina	0.991	1	0.544	18	0.761	0.863	18
Contig40_gene_762	Proteína que contiene dominio DGC	0.02	0.674	0.428	31	0.263	0.305	31
Contig40_gene_766	Dihidrorotato deshidrogenasa	0.88	0.977	0.321	38	0.569	0.716	29
Contig40_gene_769	Proteína bifuncional biosíntesis coenzima A	0	0.001	0.369	36	0.361	0.471	22
Contig40_gene_787	Subunidad H hidrogenasa B convertidora energía	0	0.669	0.296	34	0.169	0.471	22
Contig40_gene_815	proteína hipotética	0.958	0.995	0.454	18	0.402	0.821	19
Contig40_gene_824	proteína tipo adhesina	1	0.982	0.819	26	0.767	0.909	26
Contig40_gene_828	Subunidad CoxB Copaliquetinas	1	0.996	0.477	25	0.911	0.94	25
Contig40_gene_834	proteína tipo adhesina	1	0.993	0.834	26	0.674	0.853	26
Contig40_gene_835	proteína tipo adhesina	1	0.962	0.786	34	0.545	0.942	28
Contig40_gene_836	proteína tipo adhesina	1	0.999	0.593	26	0.791	0.874	26
Contig40_gene_837	proteína tipo adhesina	1	0.999	0.828	26	0.628	0.886	26
Contig40_gene_841	proteína tipo adhesina	0.903	0.992	0.461	19	0.578	0.903	19
Contig40_gene_847	proteína hipotética	0.919	0.895	0.416	25	0.548	0.877	23
Contig40_gene_848	proteína hipotética	0.965	0.99	0.498	25	0.603	0.878	23
Contig40_gene_848	proteína hipotética	0.999	0.992	0.384	26	0.531	0.872	23
Contig40_gene_867	proteína hipotética	1	0.909	0.717	25	0.426	0.85	19
Contig40_gene_872	proteína tipo adhesina	0.206	0.065	0.461	22	0.109	0.189	22
Contig40_gene_900	peptidasa señal 1	0.779	0.947	0.278	22	0.188	0.206	20
Contig40_gene_906	proteína hipotética	1	0.998	0.409	27	0.244	0.772	23
Contig40_gene_909	ribonucleasa	1	0.743	0.554	35	0.616	0.914	35
Contig40_gene_930	proteína tipo adhesina	1	0.984	0.759	26	0.862	0.891	26
Contig40_gene_964	proteína tipo adhesina	0.04	0.851	0.497	31	0.389	0.614	25
Contig40_gene_975	glycerol-3-phosphate dehydrogenase (NAD)	0.964	0.837	0.177	28	0.242	0.464	20
Contig40_gene_982	proteína hipotética	0.967	0.846	0.636	31	0.379	0.714	31
Contig40_gene_996	proteína hipotética	0.999	1	0.611	21	0.456	0.775	21
Contig40_gene_1008	proteína tipo adhesina	1	0.999	0.573	25	0.763	0.911	25
Contig40_gene_1021	proteína tipo adhesina con dominio cheilina proteasa	0.217	0.413	0.259	31	0.315	0.24	23
Contig40_gene_1025	proteína tipo adhesina	0.999	0.782	0.645	32	0.511	0.8	32
Contig40_gene_1029	proteína hipotética	1	0.99	0.495	25	0.692	0.727	24
Contig40_gene_1036	proteína hipotética	1	0.984	0.696	28	0.774	0.88	28
Contig40_gene_1074	proteína tipo adhesina	0.999	0.985	0.565	24	0.764	0.793	24
Contig40_gene_1084	proteína tipo adhesina	1	0.997	0.737	28	0.852	0.87	28
Contig40_gene_1088	proteína tipo adhesina	1	0.984	0.684	28	0.883	0.809	28
Contig40_gene_1089	proteína tipo adhesina	1	0.983	0.808	28	0.809	0.891	28
Contig40_gene_1093	proteína tipo adhesina	1	0.875	0.824	31	0.734	0.823	31
Contig40_gene_1087	proteína tipo adhesina	0.998	0.849	0.364	43	0.308	0.835	43
Contig40_gene_1098	proteína tipo adhesina	1	0.998	0.803	27	0.905	0.887	27
Contig40_gene_1099	proteína tipo adhesina	1	0.992	0.624	25	0.615	0.89	25
Contig40_gene_1104	proteína tipo adhesina	1	0.992	0.624	25	0.615	0.89	25

FIG. 6-3

Gene	Descripción	ModA	0.206	0.937	0.842	0.34	35	0.228	26	0.375	24
Contig40_gene_1106	proteína hipotética		0.991	1	0.998	0.781	28	0.46	23	0.929	23
Contig40_gene_1158	proteína tipo adhesina con dominio cisteína proteasa		1	1	0.955	0.833	28	0.802	28	0.866	28
Contig40_gene_1176	proteína tipo adhesina		1	0.999	0.975	0.527	27	0.35	22	0.761	22
Contig40_gene_1198	proteína elusorio isomerasa rnaic_ej_10rccocovina		0.995	0.999	0.978	0.57	26	0.402	26	0.559	26
Contig40_gene_1215	proteína transportadora molibdato unión al sustrato	ModA	1	1	0.987	0.762	26	0.938	26	0.944	26
Contig40_gene_1238	proteína tipo adhesina con dominio cisteína proteasa		0.642	1	0.912	0.426	28	0.477	27	0.632	28
Contig40_gene_1247	proteína hipotética		1	1	0.98	0.788	27	0.77	27	0.911	27
Contig40_gene_1254	proteína hipotética		1	0.999	0.372	0.372	24	0.24	19	0.608	19
Contig40_gene_1270	proteína ABC transporte unión al sustrato		1	1	0.982	0.649	26	0.806	26	0.835	26
Contig40_gene_1274	proteína tipo adhesina		0.005	1	0.37	0.37	24	0.136	12	0.243	20
Contig40_gene_1286	proteína hipotética		0	0.001	0.743	0.336	31	0.162	19	0.449	31
Contig40_gene_1331	proteína hipotética		1	1	0.746	0.8	25	0.775	24	0.834	25
Contig40_gene_1330	proteína tipo adhesina		0.943	1	0.984	0.532	38	0.702	38	0.854	38
Contig40_gene_1351	proteína tipo adhesina		1	1	0.999	0.86	26	0.795	26	0.932	26
Contig40_gene_1355	proteína tipo adhesina		1	1	0.998	0.86	26	0.707	25	0.894	26
Contig40_gene_1364	proteína tipo adhesina		1	1	0.998	0.896	25	0.707	25	0.894	26
Contig40_gene_1367	subunidad G tungsteno formilmetanotranso deshidrogenasa	PvdG	0.221	1	0.221	0.4	24	0.286	24	0.203	24
Contig45_gene_8	proteína hipotética conservada		0.998	1	0.576	0.272	24	0.358	18	0.432	18
Contig45_gene_21	proteína hipotética conservada		1	0.936	0.003	0.272	24	0.43	21	0.292	19
Contig45_gene_30	proteína hipotética conservada		0.379	1	0.712	0.31	21	0.43	21	0.292	19
Contig45_gene_33	proteína hipotética conservada		0.656	1	0.125	0.388	36	0.592	36	0.437	35
Contig45_gene_36	familia papitasa Cab		0.651	1	0.405	0.199	26	0.176	15	0.118	23
Contig45_gene_50	proteína hipotética conservada		0.651	1	0.146	0.423	30	0.447	30	0.368	32
Contig45_gene_54	proteína hipotética conservada		1	1	0.999	0.85	26	0.947	26	0.902	26
Contig45_gene_60	proteína hipotética conservada		0.999	1	0.061	0.216	22	0.149	24	0.653	26
Contig45_gene_64	proteína hipotética conservada		0.942	1	0.938	0.354	28	0.174	26	0.643	28
Contig45_gene_89	UDP-glucosa/GDP-mannosa deshidrogenasa		0.982	1	0.635	0.562	31	0.145	24	0.566	31
Contig45_gene_91	proteína tipo adhesina		0.999	1	0.859	0.87	28	0.529	28	0.85	28
Contig45_gene_93	proteína tipo adhesina		0	1	0.947	0.896	28	0.54	28	0.851	28
Contig45_gene_100	proteína hipotética		1	0.033	0.571	0.462	24	0.28	23	0.146	24
Contig45_gene_106	proteína hipotética		1	1	0.999	0.889	28	0.871	28	0.852	28
Contig45_gene_116	proteína hipotética conservada		1	1	0.999	0.778	26	0.742	26	0.872	26
Contig45_gene_159	homocisteína deshidrogenasa		0.086	1	0.89	0.435	39	0.152	39	0.43	39
Contig47_gene_7	proteína tipo adhesina con dominio cisteína proteasa		1	0.995	0.688	0.719	25	0.896	25	0.919	25
Contig47_gene_13	proteína hipotética		0.999	1	0.968	0.582	35	0.285	24	0.644	24
Contig47_gene_57	proteína tipo adhesina con dominio cisteína proteasa		1	1	0.998	0.819	26	0.807	26	0.834	26
Contig47_gene_60	proteína hipotética		0.006	1	0.003	0.353	29	0.225	21	0.705	23
Contig47_gene_62	proteína tipo adhesina		0.999	1	0.954	0.589	34	0.655	34	0.821	34
Contig47_gene_98	proteína tipo adhesina		0.999	1	0.986	0.086	22	0.06	15	0.405	20
Contig47_gene_125	proteína hipotética		0.999	1	0.986	0.467	26	0.176	24	0.49	30
Contig47_gene_140	proteína hipotética		1	1	0.991	0.718	23	0.731	23	0.903	23
Contig47_gene_146	proteína hipotética		1	1	0.992	0.702	26	0.887	26	0.713	26
Contig47_gene_160	proteína hipotética		0.485	1	0.821	0.817	33	0.393	33	0.472	24
Contig47_gene_197	proteína hipotética		0.942	1	0.736	0.29	30	0.256	20	0.6	19
Contig47_gene_208	proteína hipotética		0.001	1	0.738	0.451	31	0.307	23	0.65	23
Contig47_gene_253	proteína ABC transportadora permeasa		0.995	1	0.859	0.293	26	0.443	22	0.618	22

FIG. 6-4

Contig47_gene_269	.N5(M)10	0.033	0.701	0.325	23	0.177	23	0.188	19
Merilit virahidrometamoplerina	dependiente de coenzima F420p								
Contig47_gene_304	proteína tipo adhesina	0.508	0.67	0.389	17	0.686	17	0.896	17
Contig47_gene_306	familia hidrolasa plegue alifabeta	1	0.915	0.245	29	0.215	17	0.411	17
Contig47_gene_309	proteína hipotética	0.966	0.847	0.434	24	0.262	26	0.722	31
Contig47_gene_349	proteína tipo adhesina	0.008	0.319	0.116	28	0.233	22	0.608	19
Contig47_gene_353	proteína carente de dominio unión ácido nucleico OB	0.647	0.435	0.39	23	0.171	21	0.281	21
Contig47_gene_356	proteína familia deshidrogenasa reductasa cadena corta	0.936	0.889	0.357	26	0.286	26	0.171	16
Contig47_gene_375	proteína hipotética	1	1	0.614	33	0.438	33	0.783	33
Contig47_gene_381	proteína tipo adhesina	1	1	0.764	26	0.791	25	0.783	23
Contig47_gene_391	proteína hipotética	1	1	0.557	28	0.28	30	0.647	21
Contig49_gene_3	proteína hipotética	0.688	0.671	0.537	28	0.398	28	0.549	28
Contig49_gene_4	proteína hipotética conservada	0.186	0.698	0.523	28	0.398	28	0.512	28
Contig49_gene_25	proteína superfamilia transglutaminasa	0.125	0.542	0.552	28	0.463	30	0.512	28
Contig49_gene_43	proteína tipo adhesina	1	1	0.719	24	0.789	24	0.886	24
Contig49_gene_44	proteína tipo adhesina	1	1	0.694	24	0.929	24	0.907	24
Contig49_gene_96	proteína tipo adhesina	0.907	0.997	0.534	24	0.671	19	0.84	19
Contig49_gene_128	proteína tipo adhesina	0	0.106	0.431	32	0.537	25	0.208	25
Contig49_gene_152	proteína transportadora ABC unión sustrato	0	0.004	0.359	26	0.17	21	0.082	21
Contig49_gene_167	proteína tipo adhesina	0.996	0.995	0.497	25	0.24	24	0.512	21
Contig49_gene_172	proteína hipotética conservada	1	1	0.891	25	0.913	25	0.891	25
Contig49_gene_183	proteína tipo adhesina	1	1	0.19	21	0.039	14	0.179	14
Contig49_gene_184	proteína tipo adhesina	1	1	0.596	22	0.914	25	0.869	25
Contig49_gene_180	proteína hipotética	0.241	0.953	0.504	24	0.306	32	0.63	24
Contig49_gene_194	proteína hipotética secretada	1	1	0.513	31	0.731	31	0.646	24
Contig49_gene_208	Proteína transportadora ABC unión sustrato	1	1	0.885	32	0.682	26	0.744	27
Contig49_gene_226	proteína hipotética conservada secretada	1	1	0.772	26	0.918	26	0.668	27
Contig49_gene_239	proteína tipo adhesina	0.994	0.994	0.289	33	0.343	21	0.674	21
Contig49_gene_246	hipotética conservada	0.996	0.799	0.32	17	0.396	13	0.802	17
Contig55_gene_2	proteína hipotética	1	1	0.619	27	0.866	27	0.743	27
Contig55_gene_3	proteína hipotética	0.999	1	0.31	19	0.845	17	0.614	17
Contig55_gene_7	proteína tipo adhesina	0.376	0.809	0.234	33	0.11	32	0.242	25
Contig55_gene_13	proteína hipotética secretada	0.84	0.911	0.696	22	0.82	22	0.636	19
Contig55_gene_23	proteína hipotética secretada conservada	0.22	0.951	0.438	39	0.226	37	0.318	29
Contig55_gene_40	proteína hipotética secretada	0.951	0.956	0.627	17	0.741	17	0.864	17
Contig55_gene_45	proteína hipotética conservada	0.997	0.998	0.683	32	0.796	32	0.873	32
		0.999	1	0.584	21	0.866	21	0.857	21
		1	0.998	0.684	28	0.272	30	0.682	28

FIG. 7-1

Número del gen	Secuencia de Aminoácidos	SEC ID NO:
Contig40_gene_34	MVALSILLSSIAAASA	1
Contig40_gene_35	MSTFILVLIILGILLIA	2
Contig40_gene_39	MDNKKIFVIVALALLAIVAVGSVA	3
Contig40_gene_40	MNKKIFAIAAALAIIVAVGSVA	4
Contig40_gene_41	MNAKKLTIILAAALAIIVAVGSVA	5
Contig40_gene_51	MICCVLLTFSTVSAIDMDGNLT	6
Contig40_gene_54	MIAIIFEMYNVRNKR	7
Contig40_gene_63	MNKVQLSSILALVLLFLSLAVVSA	8
Contig40_gene_70	MRKEIISILVIAIIAISVIPPTAFS	9
Contig40_gene_72	MNKKREKLLAFIFIAFALINTCFILNDNLS	10
Contig40_gene_75	MMVILLITLISVPILSLTIDYS	11
Contig40_gene_105	MNINLKKITFLCLVLLGLISFNSISA	12
Contig40_gene_119	MNNQNKYSCIVLAGGMRMRMGQ	13
Contig40_gene_161	MEDRKAKFVYVWVCLAFICSTVFSMTGGLS	14
Contig40_gene_163	MNILINGTGAIGIGLGASMSIQ	15
Contig40_gene_164	MTIYITICVILLILVLEYGLEFGLTNS	16
Contig40_gene_165	MSDVGKTVITITITVTTAFG	17
Contig40_gene_169	LKSDRKAKEAFFFSIALALGLSNIAA	18
Contig40_gene_179	MINGLMDKQKVITAFGLILFLAARAFPFVWLPILGV	19
Contig40_gene_187	MFNKKMVLAISSLAVIFASMCIVSA	20
Contig40_gene_203	MKTNLKKTIIILALLMALLISIGAIISA	21
Contig40_gene_228	MNSKGYLVLFLLIILSFSIISASEFA	22
Contig40_gene_231	MKNLNLKNIILILSLIFLVLISGSSEFA	23
Contig40_gene_232	MKRNIYFIILLVTLFLISMVVSVA	24
Contig40_gene_248	MKKMEMASYIILIASVLAIIYA	25
Contig40_gene_251	MPKIAKLWNLADPKNIPRLFAVI	26
Contig40_gene_252	MFNLAIWVYLGALAIIFGSLATWVGPGVKDF	27
Contig40_gene_260	LFAIIVLSAVSASDDFSSSLADSDSD	28
Contig40_gene_269	MKRRYKVLFLAIIITISINAISA	29
Contig40_gene_296	MLFSVIATVVSATC	30

FIG. 7-2

Contig40_gene_297	MFKIRRDTLIILL	31
Contig40_gene_306	MKAVIPAAAGLGRFLPATKRPQPKEMLPVYDKPTIQ	32
Contig40_gene_317	MIKTDVLVIGAGPAGSAAARFAAKGG	33
Contig40_gene_346	MEDNLLKNRKLLISIFLVSLLAISAVSA	34
Contig40_gene_349	MNRNKIIYLLVLLIAVVGFTMPACA	35
Contig40_gene_352	MKKSFKLIIALAILLAVSIVSS	36
Contig40_gene_411	MKKNIFLLAIIILAVVAVSGC	37
Contig40_gene_431	MLALLGLSAVAADVADPLTDN	38
Contig40_gene_448	MSENNRPLITIGIGAFII	39
Contig40_gene_466	MGKIFKIVTIIIVIALIILGVFIYSDGHSE	40
Contig40_gene_483	MDKKTIIIAAVAILVI	41
Contig40_gene_501	MKLNKFFIISIIILIFLSISAISA	42
Contig40_gene_553	MKKKIALILGIALLAFLVIGASSA	43
Contig40_gene_636	MNFNKILLIIALVFIASVGIVAA	44
Contig40_gene_721	MKRSIIFLTIILSLFLVIGYASA	45
Contig40_gene_730	VGITFTAITGALGGTTFSEPLGNVLSQFIPY	46
Contig40_gene_732	MDSKKLIIVTALAFIAVSTASVSA	47
Contig40_gene_733	MNVNKKIFLLVIFLISISIAGVYCADLHQDSDLTA	48
Contig40_gene_749	MILALFCEFIVIGSASAADFK	49
Contig40_gene_750	MISLLIISILISAASAA	50
Contig40_gene_762	MEEKIALAACSGMSPNGLVARVAVHDLAID	51
Contig40_gene_766	MLKTKLGGISLKNFIMLAAGVLGSHA	52
Contig40_gene_769	MEIVLCVTSVAAVETVKLAREFKRQG	53
Contig40_gene_787	MVVAIIIFASSLFDALYGFRN	54
Contig40_gene_815	MILAILLAVGMTLAVSA	55
Contig40_gene_824	MNKRIFLYIALIFIIISLISFSAVSA	56
Contig40_gene_828	MRYNKKIFFELFLCLLIPOAIY	57
Contig40_gene_834	MNSNRTYAVLGLLIIILSISGAISA	58
Contig40_gene_835	MINKRIISLSLIIILVFLIIGLSAVSA	59
Contig40_gene_836	MDFKKAIPLFALLLIIILFIIGSSSA	60
Contig40_gene_837	MKLKKSFTLAVLIVAILAIGAVSA	61
Contig40_gene_841	MILISLIIILVILSISCVSA	62

FIG. 7-3

Contig40_gene_297	MFLKIRRDTLIILL	31
Contig40_gene_306	MKRAVIPAAGLGRFLPATKAQFKEMLPVYDKFTIQ	32
Contig40_gene_317	MIKTDVLVIGAGPAGSSAARFAKGG	33
Contig40_gene_346	MEDNLLKNRKLLISIFLVSLLAISAVSA	34
Contig40_gene_349	MNRANKIIVLLVLLIIVVGTMTGEPACA	35
Contig40_gene_352	MKKSFKLLIALALILLAVSISS	36
Contig40_gene_411	MKKNIFLAIILLAVAVSGC	37
Contig40_gene_431	MLIALIGLSAVAADADPLTDN	38
Contig40_gene_448	MSENNRLLTIGIGAFII	39
Contig40_gene_466	MGKIFKIVTIIIVIALAILLVFIYSDGHSE	40
Contig40_gene_483	MDKKTIIIAAVAILVI	41
Contig40_gene_501	MKLNKFFIISIIILIFLSISAISA	42
Contig40_gene_553	MKVKIAILIGIALLAFLVIGASSA	43
Contig40_gene_636	MNFNKKILLIILVFIASVIVAA	44
Contig40_gene_721	MKRSIIFLIIISLFIYGYASA	45
Contig40_gene_730	VGITFTAITGTALGGTFSEPLGNLYSQFIPY	46
Contig40_gene_732	MDSKKLIIVTALAFIAVSIASVSA	47
Contig40_gene_733	MNVNKKIFLLVIFIIISISIAGYICADIIHQDSDLTA	48
Contig40_gene_749	MILALFCFIVIGSASADFK	49
Contig40_gene_750	MISILLIISILAIASAASA	50
Contig40_gene_762	MEKIALAACSGMSPNGLVARVAVHDLAID	51
Contig40_gene_766	MLKTKLGGISLKNPLMLAAGVLGSHA	52
Contig40_gene_769	MEIVLCVTGSAVAEIVKLAREFKRQG	53
Contig40_gene_787	MVVATIIFASSLFDALYGFKN	54
Contig40_gene_815	MILAILAVGMTLTAUSA	55
Contig40_gene_824	MNKRIFLYALIFIIISLIFSAYSVA	56
Contig40_gene_828	MKYNKKIFELFLLICLIIPQAIY	57
Contig40_gene_834	MNSNKTYAVIGLLELLIISIGAIISA	58
Contig40_gene_835	MINKRIISLLELLIVFLIIGLSAVSA	59
Contig40_gene_836	MDFKKAIPLFALLLILFIIGSSSA	60
Contig40_gene_837	MKLKKSIVILAVLIVAILAIGAVSA	61
Contig40_gene_841	MILISLILVILSICVSA	62

FIG. 7-4

Contig40_gene_1247	MNYSIIIFIFIMDALVLMASIQVCGA	95
Contig40_gene_1254	MKFNRSVTLGILSLLFVTLTVSSVGA	96
Contig40_gene_1270	MEKFTTIIIVTIIALIALAC	97
Contig40_gene_1274	MKNKSLIILISLIIITIIISIGSVVA	98
Contig40_gene_1296	MRSSTILLSASTAESRSPSLTTR	99
Contig40_gene_1331	VLLICFIGLVEAIIIMALVDWEDLAI SVRKSP	100
Contig40_gene_1350	MNKKIILSLLVLLVAISVAVAA	101
Contig40_gene_1355	MSLSIFVLVIGGGFINKRILII FVLIFFISIGSVVA	102
Contig40_gene_1364	MNKKIIMSFLVLLJLAISVAVSA	103
Contig40_gene_1367	MKIQRGIYIILTLVLFSLSAASAA	104
Contig45_gene_8	MELKYDDKCLGCGVCVIACPVNASISFEN	105
Contig45_gene_20	MNRRSKLIILAIIVII	106
Contig45_gene_21	MKRSKLIIAILVLLGLL	107
Contig45_gene_30	MKEYKTAIIGGGFAGMIAIRAABIILGFNAVCILE	108
Contig45_gene_35	MANEGGHLKTIIMIIILAFICGL	109
Contig45_gene_36	MDNKIKAGIALAIIVLVAVIGSFINESN	110
Contig45_gene_60	MPKVSILIVCLVSLFLLVQASA	111
Contig45_gene_64	MWYDMKRRRFYLIIFLIIILAAIA	112
Contig45_gene_89	MKITVAGVGVGLSLAVLLAQKHDTA	113
Contig45_gene_91	MNLMKITVAGVGVYGLSIAIILLAQKHDTA	114
Contig45_gene_93	MEIRYKNLLKVTTFELVLLISCGFASA	115
Contig45_gene_100	MKIRYKNLLKVTTFELVLLISCGFASA	116
Contig45_gene_106	MQRSLFDKVRTSLWMLPSPFGLV	120
Contig45_gene_116	MKFKNSHILLVSLISIFLLISIAASA	121
Contig45_gene_159	MNSKKTAVLGIILLSFAIVGSASA	122
Contig47_gene_7	MDECKLVLIGFVAVGQVARAISMKKEMINERFGISLKV	123
Contig47_gene_13	MNRKTIIFGVIFIVFILFSISTVSA	124
Contig47_gene_57	MNKQNVFALLIITLISVVAVS	125
Contig47_gene_60	MLNKKIILITFILISASSASASA	126
Contig47_gene_62	MGLVASVAGGIFFEAGMIATCT	127
Contig47_gene_98	VFSVSLNKLKIGRVFICLFLVLFISCSINCVFA	128
	MGGEIINNEKLLIILIT	129

FIG. 7-5

Contig47_gene_125	MDKMTVSVAFLLLLLAVALSVFDESNS	130
Contig47_gene_140	MKLSRIIVLMLLIFFAGMVYA	131
Contig47_gene_146	MDSKKIIVIGLITVLAIFLASSVA	132
Contig47_gene_160	MKPYVILIGSASGIGKSTVAELAKTLNIRHLV	133
Contig47_gene_197	MLISVLGVIIIMVVAA	134
Contig47_gene_208	MATRTKQTCIRLSYFHGREFL	135
Contig47_gene_253	MELSKSDKYLIWVGIIFFCLAL	136
Contig47_gene_269	MKVAIIGAGCYRTHAASGITNF	137
Contig47_gene_304	LSVILILFLAVSTVAA	138
Contig47_gene_306	MNKKLKIILYILLALI	139
contig47_gene_309	METKNLIIICATVILLAVIVLSAFIYVNMGN	140
Contig47_gene_349	VIVLAFAIIFLGYISLGNNO	141
Contig47_gene_353	MELNDEIIFKVALITALVGMIG	142
Contig47_gene_356	MKNYEDIKDKAVVVGASSLGGWQI	143
Contig47_gene_375	LTFNNLRINIKDCMVFVVFVLLLSILAVSA	144
Contig47_gene_381	MKFKKYLFIILLIALICIIISVAVA	145
Contig47_gene_391	MDKMTVLLVALFCLLCVGS	146
Contig49_gene_3	MDRKDIIIIILVLIISLLALGLRHHQV	147
Contig49_gene_4	MTSEIMILPTAVVLAADSAVTISDIK	148
Contig49_gene_25	MRKKILFTLMILICFTLNSVCA	149
Contig49_gene_43	MRLRYFAIISLILLIFLVPVSEFA	150
Contig49_gene_44	MFIGLLIIGLLIIPISFA	151
Contig49_gene_96	LLIGLVICAGVYFQFNYPYPTYL	152
Contig49_gene_128	VVLVAVVIGSTAFLELLNYDETVKYT	153
Contig49_gene_152	MDKKTALIAIIVIALVAVG	154
Contig49_gene_167	MNNKTLFIIGLFIKLLFTIIPWVA	155
Contig49_gene_172	MIKTDNKGQITVELLLLSF	156
Contig49_gene_175	MLNRKALIFSLIVLFMLSISAVSA	157
Contig49_gene_180	MDNKAIIGIVIALIVIVLACFAY	158
Contig49_gene_183	MSYFNKGHINWILLICLLIGTLAMGSASA	159
Contig49_gene_184	MKNKAMFLISALLIIVILSISAVSAARDDATA	160
Contig49_gene_194	MKFNKSLIAIFVILLIVAFSSISVIA	161

FIG. 7-6

Contig49_gene_208	MDKKIIGAVVALLVIIVGA	162
Contig49_gene_226	VEGDNNVNIKTVALAV	163
Contig49_gene_239	MLIMAILVLLTMAVSASELEDIQVT	164
Contig49_gene_246	MNNTTKILIGVLMGLL	165
Contig55_gene_2	METENLIIVILLVLIAMAGIF	16
Contig55_gene_4	MELVILLFAFIVIGGSYSVFA	167
Contig55_gene_7	MALLIAMSCVSAGNASDNLDDLTISDSNSLDIVSTSN	168
Contig55_gene_13	MNNKYFLGIITIIIVAV	169
Contig55_gene_23	MLNDKSELKSLFLFLVLIITSFNSVYA	170
Contig55_gene_40	MKKIILGTCLIFLLISVAYA	171
Contig55_gene_45	MKINLKRVIILGIIILILICISSASIISA	172

FIG. 8-1

Número de gen	Secuencia Nucleotídica	SEC ID NO:
Contig40_gene_34	atggctccttgcccttaagcacaatcctactcagttccaattgacagcagcatctgca	173
Contig40_gene_35	atgagtacttttatttagtcataaataacttgggtattttaaataatcgca	174
Contig40_gene_39	atggataataaaaaataatttggttatttggcttttggctcttttagctatcgttggcagtaggatcctgtag	175
Contig40_gene_40	atgacaataaaaaataatttggccattgacagcttttagccattatagctatcgtagcagtaggatcagtcagtg ct	176
Contig40_gene_41	atgaatgctaaaaaactaatcttctagcagccttagctattctcgtctatctcgtctatcgttgcagtaggctcagtaagtg ct	177
Contig40_gene_51	atgatttgcctggtttttattaacattttcaactgttagtgcattgatggatggaaatctaaact	178
Contig40_gene_54	atgattatggcacaatcctcatatataaagggtccgcaataagaga	179
Contig40_gene_63	atgaataaggttcaabtgctcctccataacttggcttttagtataataatttcttggcttggctgtagttagtg ca	180
Contig40_gene_70	atgagaaaaaataatttctatatttggtaattgctattatagcaatctcagttattccaactgccccttttca	181
Contig40_gene_72	atgaataaaaaaagattttaaattatttaactattttatagcatttggcaccattaaacacttggtttattt	182
Contig40_gene_75	taaatgataatctctca	183
Contig40_gene_105	atgatggtcattctactaataaacactcctttctgttctctatcctcactaacaattgattatcca atgaatatttaatttaaaaaaatcacttcttatggttttagttctcaatcgggttggatctcatttaatt caatcagcgcct	184
Contig40_gene_119	atgaaataatcaaaaataagttatcttgcagatgttttagctggagcagtagcagaagaatgggtcag	185
Contig40_gene_161	atggaagatagaaaaagcaaaatttatcgtttatgctgttgtatcgttacttggcttttactgcagcagcag tcttctctatgactggcgtcttct	186
Contig40_gene_163	atgacaataactaatcaatggaaactggagctatcggaaatagggttggagcactatgatttcacaaggt	187
Contig40_gene_164	atgataatagtcactacaatctgtgttatcttaatttggatagttctttttatggatbgttccctggattga caaacagc	188
Contig40_gene_165	atgtctgatgttggtaaaactgtaataacaactattattactcttagtaactactgcatitgggt	189
Contig40_gene_169	ttgaaatcagataaacgggctaaatttggccataatcttctcaattgcaatccttggccttgggactggagcaata ttgcagct	190
Contig40_gene_179	atgattaatggaataatggacaagcagaaggttataactgcctttggcataattctatttttggcagctgctt tcagtcctgttggtagctcttgcctatcttaggagt	191
Contig40_gene_187	atgtttaaataagaagatggttttagccataagcttattagctgttatcttggcatctatgtgcatagtttcag ca	192
Contig40_gene_203	atgaaaaaaaatctttaaaaaaacaacaatcatalttggcactgctgatggccattttaaatttatcagattggag ccatctctgca	193

FIG. 8-2

Contig40_gene_228	atgaattccaagggaaaaatactctgttttattcttcttaataattatcatttagpataatctctgttcat ttgct	194
Contig40_gene_231	atgaagaaaaattaaagttaaaaabattttaattttatcaattaatctctctttttgattaaagcataggat cttcatttgca	195
Contig40_gene_232	atgaaaagggaatattttattttattttttatttagttacacactatttttaactcagtagtggtagtgca	196
Contig40_gene_248	atgaaaaaaatggaaatggtagttatattatcttaattgatctgtattagctatactttatgca	197
Contig40_gene_251	atgctaaaaattgcaaaatggaataagtagcagatccaaagaacattcctaggtgtttgctgtaatt	198
Contig40_gene_252	atgtttaacttggctattttgggtttatttaggtttggcattagctatttttggaaagcctcgaactgtatggg gtcctggagtaaaaggatcca	199
Contig40_gene_260	ttgttcgctatagtaagcctatctgagtcagcgaagcagatgatttttcaagttcccttgc	200
Contig40_gene_269	atgaaaagaagataaaaagtttttctatttggccatcttaactataaagaacattaatgcccatttcagct	201
Contig40_gene_296	atgctcttttcagtaattgtctactgtatctbgtactctgt	202
Contig40_gene_297	atgtttattaaaaattagaagagacacttttaataatattat	203
Contig40_gene_306	atgaaagcagtcattctcagcagggcttggaaacaagattccttccctgactaaagctcaacccaaaagaga tgttgccggtttatgacaagccgacattcaa	204
Contig40_gene_317	atgattaaaaactgatgtatfgttatttggctggaacctgctgggtcttcaagctgtagattgcaagctaaaag gcggc	205
Contig40_gene_346	atggaggataatcttttgaaaaatagaaaaactaaatttgataagtatcttccctttagtctgttgcgaattt ctgctgtaagcgca	206
Contig40_gene_349	atgaacagaaaaataaaataattgttttgccttatttattgtagcagttgttggctttacaatggggccagctt gtgca	207
Contig40_gene_352	atgaaaaaatcagtttttaaaattcctaaattgcttttagctttaattttatttggctgtatcaattgttccatct	208
Contig40_gene_411	atgaaaaagaatatttttttaattgcaataataactaaattgcagttgttgcagttagtggtgt	209
Contig40_gene_431	atgttgatggccttacttggcttctctgctgttgacgagcttgacgctgacccatcaactgataat	210
Contig40_gene_448	atgagcgaaaaaatagaaactttgattacaaataggaatcggcgtttttattata	211
Contig40_gene_466	atgggaaagatatattaaaaattgttacaatcatalatgattgtcattggcctatacttgggttttccatct attctgatggacattctgaa	212
Contig40_gene_483	atggataagaaaaacaatcattatagctgcagtagtatctctcgttatt	213
Contig40_gene_501	atgaaattaaataaattcttcattatcagcacaataattgattatattttctatcaatagtgcaataagtcg	214
Contig40_gene_553	atgaaagaaaaaatagcaattattttagaaattgcattattagaacttcttagtcaatcggcgcatccagcgca	215
Contig40_gene_636	atgaatttcaataaaaaattttattaaatagcatttggatttcaatcttggctagtttggcaattgttgcagct	216
Contig40_gene_721	atgaaaaagatcaatcatatttttaacaattatattatccttatttttagtaattggctatgcaagcgct	217

FIG. 8-3

Contig40_gene_730	gtggcataaacctttacagcaatcatcacaggggcaattagtggaactacttttcagaaaccttttaggaaact accttagccaattcataacctaac	218
Contig40_gene_732	atggattctataaaaaactgatttttagtgactgcatgggttttttagctattgtttccatagcttcagttagt catgggacttggttggaa	219
Contig40_gene_733	atgaatgtgaataaagaaaaatattttacttgtaatctttataaatctatttcaatagctggagatattgtg cagacatccatcagatagcgatttaaccgca	220
Contig40_gene_749	atgatactggcaactatttggttttatagtcattggctcagcaagtgcagcaacttataa	221
Contig40_gene_750	atgatctcactgcttcttatttcaatctctgctataagcgcaagctgctgca	222
Contig40_gene_762	atggaagaaaaaattgcttttagctgttgcttgaagtgatgagtcocaaatggtttggtygcaaggttcagttc atgatttggctattgacg	223
Contig40_gene_766	atggttaaaaaactaaaattatgcyggaattagtttaaaaaaccattaatgcttgcaggtgttttgggaaqcc atgca	224
Contig40_gene_769	atggaattgtattatgtgaaacagcgagtgtagctgcagtgaaactgtaagttagctcgtgaatttaagc gtcaaggg	225
Contig40_gene_787	atggtagttgcaacaataatcttgcatccagcttattgacgcccctttatggatttaaaaaac	226
Contig40_gene_815	atgatattggcaatattgcttgcggttggaatgacacttactcagtaagtcca	227
Contig40_gene_824	atgaaataagcgaatatttctatatatagcaactgattttattatttccctgcttttcttctgcagtcagtg ct	228
Contig40_gene_828	atgaaatataaaaaagatattcttttattttattattgtgtcacaataatcccaagctattat	229
Contig40_gene_834	atgaaattctaataagacttgcagattagggattatgctcttataatctatccatagggcgtatttagtg ca	230
Contig40_gene_835	atgataaataaaaaaataatagctttagctctgctgattatattggtcttctcattattggattaaagtgcag tcagtgct	231
Contig40_gene_836	atggatttttaaaaaagcaatccctctatttgccttattgctattgattttattattatcctcttcaagcg ca	232
Contig40_gene_837	atgaaagcttaaaaaagtttccagtcatttttagcgtatttgcctttagcaataacttgcattggggtgtaagtg ca	233
Contig40_gene_841	atgattttatcccttaatttttagttatcttaagcatttcttgtgtaagtgct	234
Contig40_gene_847	atggcaactccaatataactcagtaaatatagattattgatttgcagcaggagtaactgca	235
Contig40_gene_848	atggaataatcagaactcttataatccgtaactcactgctttatgatttgcagcaggagtaactgccc	236
Contig40_gene_867	atgagaagaagaaattttaaattgcagctattgcaatcattataatattatgtggaggtgatttggca	237
Contig40_gene_872	atggtgatatacaatlgacttatctctcattctttagctttaggtgcaagtgca	238
Contig40_gene_900	atgaaagaaatfgctattttatctcactcttataataaattgttcttatttgcgcacaacactta	239

FIG. 8-4

Contig40 gene_906	tigtgtgaagcaggatgatgctcttccactcgggttggctggactgccttgctgtggggctt	240
Contig40 gene_909	atgaaaactggaaaataatggatttaataataatcaatcctcttgctgtcgtttcaagttagggc	241
Contig40 gene_930	atggaaaataaaagatttccatttttacttttaagatggtgcagctattatgcgttcgagctttcagcaaa	242
	atgactgtggataaacttggaagttagat	243
Contig40 gene_964	atgagtttaaacgaaataattacttgcaggttaatactgtattataataatatttcaatttctgtaagtg	243
Contig40 gene_975	caaatgaaaa	244
Contig40 gene_982	atggataaaggtaggaattatagagcaggtagctctaggtacagcttttagctcaaacagtgggcta	245
Contig40 gene_996	ttgactctgtagcatcacggcctctggcctcactgcactgcagctctatgtagggcct	245
	atgaaaatacaagatataccttatatttcttcttttggttatttttggaaaataggacctgttcacgtcat	246
Contig40 gene_1008	ataccatgtaaatgct	246
	atgatttttaatttctcattctcttattcactatttcattactcgtctatoggtcgcgaaaggcga	247
Contig40 gene_1021	atgaaaattataaaaaatagcacaatcatttttttataataatttcaatttctatogattggagcagctgcagct	248
Contig40 gene_1025	ttggcagtgatttggataactctatttccacttgggactgttgcagcagtgaaaaatatagtttatt	249
Contig40 gene_1029	atggaaaaccctaaaagtattatggaagcagttatttggatttctttagggccttattgtctgtatttcta	250
	tagtttccacctatagcagct	250
Contig40 gene_1036	atgaaataaataaaagatttggccggatagcacaattggctatttggctatttcttaattgggacagttgct	251
	atggagaaaactatgaaatctaaaacttttatacttctaatactattctctatttctaataageatttccatcag	252
Contig40 .gene_1074	tttcagca	252
Contig40 gene_1084	atggataaaagatttttataagtttagctttattctgctagctattttcacaatagggcgttggcgct	253
Contig40 gene_1088	atgatgaaaaatgactaaaagaattcttttttaataattagtttaataactaataatttcttacaatttggctg	254
	tcagcgca	254
Contig40 gene_1089	atggtgattatgaataataaaaagcttttattgttagttgattatacctaactatttgacaaataggcgtg	255
	tcagtgca	255
Contig40 gene_1093	atgaaagtttaataaaaaatagggcatactgccataatcaataatttataatttctattttaaagtattctatgg	256
	catctgct	256
Contig40 gene_1097	gtgttttattttgaaatttgaaataaaaagtttataataatcatttcatttcataatattgatatcttatcta	257
	ttggaaatggcactgaa	257
Contig40 gene_1098	atgcaagcaattattcaccagttaaaagacaatttctcatttttagtgacaataatgaagaaaagttatttaaac	258
	gtatattcatatgttttagtctcttctcttactgtcttctgattgttggcagtaagtct	258
Contig40 gene_1099	atgaaactttaaaaaacttttaagtatttcatttaacactttatttatttgtcttatacagtaggttttagcagcaaa	259
	ggcct	259
Contig40 gene_1104	atgaaaataaagaggtttgctatatttcttaattatctctgatttttttcttatttctactatctgcaggttgcagct	260
Contig40 gene_1106	gtgactgtttcagttttttataagtgcttcaatttggcttttgcttttggcaatgttcttaagcaatgttcagataacgga	261

FIG. 8-5

Contig40 gene 1158	atgaaggctcttaaaagatagcaattatcatgcttattttaatacatatctctggaggggtttcagca	262
Contig40 gene 1176	atgaatttttaaaacaaaggaagcttgattctctatttcattacttttcatttaataataggtatggaaatgg catcagca	263
Contig40 gene 1198	atgggaaatttaaaattatatttattctagtttttagctctatttttaaatatgtggaattgct	264
Contig40 gene 1215	atggattctaaagaaaaatattaatgatggtctgagttgctttbaatagcaattgctgtgaagttcatgctctg ca	265
Contig40 gene 1238	atgaagttaaaatcaaagtaTTTTGtatttttaactcataaataatgatacctattccaacagtttcag cg	266
Contig40 gene 1247	atgaattattccattattatcttctattcttctatggatgcatgggttaattggctagcacaagctct gtggagct	267
Contig40 gene 1254	atgaagttaaaatcaaagtttttaggattttatctctatttattttttcaaatcttgtttccaagtgg gggca	268
Contig40 gene 1270	atggaaaagaaaactcaaatatattggttatttttaattgctctttatgcatgc	269
Contig40 gene 1274	atgaagataagagtttaataattttcttatttattactgattacaataataagcatalggtctgtttg ca	270
Contig40 gene 1296	atgagaagcactatcctggttaagtcaagtactcggaaaagccgtcaccagctcacaacagggcaga gtgcttctcatttggctttataggattggttgaggcgatactgatggcaattggttgattggaggacttggcaa	271
Contig40 gene 1331	tatccgttcgcaagctcct	272
Contig40 gene 1350	atgaataaaaaaattatcttaccctccttttagtattatagtagctatttctgctcagttgcagca atgctgattatccatatttctggtttatagggatttataataaagaatatttataatattcgtat	273
Contig40 gene 1351	ttctaatattcttcataagtatggatctgtagttgc	274
Contig40 gene 1355	atgaacaataaaaagattataatgtcttttctattggtcctatlgattgcaatatctgctctgcagtttcag ca	275
Contig40 gene 1364	atgaaaatccaaaagggtatatataataataaacttactgttctcttagcccttctgctgcaagcgcag ca	276
Contig40 gene 1367	atggaaacttaagtagatcaagataaaatggttaggttggaagtatggttatcgcatgctgtaaacgctt ccatcagtcggaaaaac	277
Contig45 gene 8	atgaatcgaagatcaaagtttaaatgcaatttttaatagttatcataata	278
Contig45 gene 20	atgaaaagatcaaaaaaataataatagcaatctctgttgaatccttttgggatacta	279
Contig45 gene 21	atgaaagaaatacaagatagcaattatagggagggccaggaatgatagctgcaataagagcgcagaaaa tattagggcccaaatgcagtagcatctagag	280
Contig45 gene 30	atgggcaaatgaaggtggaggacaatttaaaagactatttttagattataaataagctttatttgggtgc tt	281
Contig45 gene 35	atggataataaaatcaaagcaggcaattgcaacttgcaataatcgttttagtggctgctgcttttcattca	282

FIG. 8-6

	ttaatgaaagcaat	
Contig45 gene 36	atggtttaaagtaagcaaaaagcataattaatcgttggtagtacccttcccttccctatttagttcccaagctagcg	283
	ct	
Contig45 gene 60	atggtggatgatatgaaaagaagaaggttttataataattataaattctatttgatttagoagctatag	284
	ca	
Contig45 gene 64	atgaaaaattacagttgcgggtgtaggatatgtagggtttccacttgcctgctgctcaaaaaacatgatg	285
	ttacagct	
Contig45 gene 89	atgaaatttgatgaaaaattacagttgcgggagtgagatagtgaaggcttctattgctattctgcttgcocaga	286
	aacatgatgtaaccgca	
Contig45 gene 91	atggagataagataataaaaatttataaaagtttttactatttttctcttcttcttactatcagttgcggatttg	287
	cctcagca	
Contig45 gene 93	atgaaataagataataaaaatttataaaagtttttactattttctctgcttactatcagttgcggatttg	288
	cctcagca	
Contig45 gene 100	atgcaacgcttcattttatgataaaagttaaacaacoccttatggatgcttccatccttttttggatttgta	289
Contig45 gene 106	atgaaatttaagaaattccacatcttactcgtttccatttaataatccatcttctattgttgagcataaagcag	290
	cctcagca	
Contig45 gene 116	atgaaattctaaaaagatagcaattgttctctggaaataatgttcttccatttgcatttgcattgaggctcgtcag	291
	ct	
Contig45 gene 159	atggatgaatgtaaaccttgatataactcggttttggcgtgtaggccaaaggtgttgcacgtgcaatccatga	292
	aaaaggaaatgatcaatgagaggttggcataagcciaaaagta	
Contig47 gene 7	atgatgagaaacaacatatttggagttatatttctcgttttttatttctcagcatttcaacggtttccagca	293
Contig47 gene 13	atgaaatacaaaaacgctatttgccttgatattataaacaacatcttctttatcctgtagtgcgtcagc	294
	atgttgaataaaaaataataaatttcaacatttttataatattgctatttctcagcaagtcgcatctg	295
	ca	
Contig47 gene 60	atgggagttattagctagttgctgagggcatttttttgaagcaggcagatgattgctacttgaca	296
	gtgttttcagtgagtttaataaataaaaccttaagattgtagagtttttatttgccttttatttagttttattt	297
	cctgttcaattaattgtgttttgca	
Contig47 gene 98	atgggaggtgaaataataaaatgaaaaataaagtttaattttaaattttaacc	298
	atgataagaaaaatgattgttccagtggttttcttttatttgatttttggcagtggtctctctgtatttg	299
	atgaaaagcaatagc	
Contig47 gene 125	atgaaaatctcaagaaataatgtattattaatgattctaatcttcacgcaggaatggtttatgca	300
Contig47 gene 140	atggattctaaaaaaatattgatttttttaggtttaaactgttttagctatttttttagctagttcagttagtg	301
	ct	
Contig47 gene 160	atgaaaaccatagttaattctcataggaagccttcagggatagaaaaatccacagttgcagctgaacttgaaa	302

FIG. 8-7

	aaacattaaacattaaagcacttggtg	
Contig47 gene 197	atgtaaatatcagctctbggagtgattgtaattatcaattatgtagttgcagct	303
Contig47 gene 208	atggcaaccagacaacaacaactatatacaggttgattctccatccatggggctgttcttcta	304
Contig47 gene 253	atggaattaaagtaaaagtgacaataatattaactgtagtaggattatattctctctgcatca	305
Contig47 gene 269	atgaaagttagcaatttttagtctgctgttacagaacctcagcagctagtggaattacaataat	306
Contig47 gene 304	ttgtcagttattctgattctgttttttagcagtttcaacggtagctgca	307
Contig47 gene 306	atgataagaaccttaaaataatccctttatattttatttggctttaata	308
Contig47 gene 309	atggaaactaaaaatcttataatcatttggcaacctgtaatatggcgttgtaaatagttttatcgtcttcta	309
Contig47 gene 349	ttatgtcaacctgggcaat	
Contig47 gene 349	gtcctgttttagctttttgcaataataatattctagatattccatcctctaggaaaaaatca	310
Contig47 gene 353	atggaattgaaatgatgaaataataattaaagtggcaactgattactgctgggaatgattggg	311
Contig47 gene 356	atgaaaaatttttcgacataaaagcaaaagtagcagttgtaaccggcttctccggattaggttggcaaa	312
Contig47 gene 375	tt	
Contig47 gene 375	ttgaccttcaacaaccttagaataaacaattaaagattgcattgcaattttgtagtgtttacagattgcttt	313
Contig47 gene 381	tatctattttagctgtaagtc	
Contig47 gene 391	atgaaattcaaaaaatatttatttctctgctaataagcttatttgcataatcaagtgctctgctgagttgctg	314
Contig49 gene 3	atggataaaaaaatgacagttttatttggttgcccatttttggccttctctctgtgtaggctca	315
Contig49 gene 4	atgataaaaaagacataataatcactcgtttctcataatacactatcactattggcattgggcttcaata	316
Contig49 gene 25	atcaatcaagt	
Contig49 gene 25	atgacatctgagattatgattttaaacccaactgcagtggttttagcggggaagtcagttacaataagcg	317
Contig49 gene 43	atataaaa	
Contig49 gene 43	atgaggaaaaaagatccttttccctaactttgatgatgataactatctgttttactttaaacagcgtttgtgct	318
Contig49 gene 44	atgagattaaagatattttgcaataatagtttaattcttcttataatttttagttccagttagtttgca	319
Contig49 gene 96	atgtttattggcttattataataggtctattataatcaatccctataagttttgct	320
Contig49 gene 128	ttgctaattggacttgctcactctgtagggtctttttatttccaatteaaactgcaactccacatatct	321
Contig49 gene 152	ct	
Contig49 gene 152	gtggttttagttgctgttagtgattggctctactgcattcctattaaattatgatgaaactgtaaaataca	322
Contig49 gene 167	atggataaaaaaactctagcaattatgctattatcgttatagctctttagctgttgggt	323
Contig49 gene 172	atgataaaaaacagcaataaaagacagataaacagtcgaactgctcctcttttaagcttt	324
Contig49 gene 175	atgataaaaaacagcaataaaagacagataaacagtcgaactgctcctcttttaagcttt	325
Contig49 gene 180	atgttaaatagaaaggctttgatttttccattgtgttttattatgctatccattctctgctgttccagct	326
Contig49 gene 180	atggataaaaaagcgaataatggaaattgtaattgcaattgcaattgcttcttctgctttgctttat	327

FIG. 8-8

Contig49 gene 183	atgagttattttaaaaggacatatatggaatattttatlaaatttgcttctcatcggaactttggctatga tgggttcagcaagtgcc	328
Contig49 gene 184	atgaaaaacaaggccaatgTTTTAAATctgcattattgatagcagttattctatctctcagtgctgtaagtg ctgcagatgatgctattgct	329
Contig49 gene 194	atgaaatttaacaagaggtttaattgcaatttttgtaattttgattgttgctttcagttcccatatctgtcattg ca	330
Contig49 gene 208	atggataaaaaaattattatcgggtgagtggttgcaactcttgttataaattgttggtgct	331
Contig49 gene 226	gtggaaggtgataaataatggtaaatataaaaaactgttgcattagctgtt	332
Contig49 gene 239	atgctgattatggcaattcttctgttttattgacctggccaggttaagtgccagcgaacttgaagacattccaag tcaca	333
Contig49 gene 246	atgaataacaactactaaaaatlaattggagttctctatgggaactgctt	334
Contig55 gene 2	atggaacacagaaaaatttaattattgtaattctttttagtttaaatagctatggccggaaattttc	335
Contig55 gene 4	atgTTTTTGGTTATATbattatttgcaatttatgttataaggaggatcttattcgggtatttggct	336
Contig55 gene 7	atggcTTTgcttattcttgcaatgtcatgtgtgtctgcaagcaatgcaagtataaatttggatgatttaacca tttcagcagtaattcactagatcttgtatcacatcaaat	337
Contig55 gene 13	atgaacaataaaactcttttttagggaataataataataaataatgycagtt	338
Contig55 gene 23	atgcttttaaatgataaaactcgaactattaaatcattatctatttttttggctaatagttctaattacaa gttttaattcagtttatgca	339
Contig55 gene 40	atgaaaaataaattcttggacaatgtatctattcttattcttggattagtgctgcacatagca	340
Contig55 gene 45	atgaaaaataaatttaaaaagagttatttttgggaattatttttgatttttggattttccctcagcaagtatca tttcagca	341

FIG. 9-1

Número de gen	Secuencia nucleotídica de codones optimizados	SEC ID NO:
Contig40 gene 34	atgggtgctggcgtgagcattatctgtcgtgagcagcattgcgcgcgagcgcgcg	342
Contig40 gene 35	atgagcacccttattctgtgtgattattattctgtggcattattctgattattg	343
Contig40 gene 39	atgatacaaaaaaattttgtgattgtggcctggcctgctggcgattgtggcgggtggcagcgtgagcgcg	344
Contig40 gene 40	atgaacaacaaaaaattttggcgattggcgctggcgattattgcgattgtggcgggtggcagcgtgagcgcg	345
Contig40 gene 41	atgaacgcgaaaaaactgaccattctggcgcctggcgattctggcgattgtggcgattgtggcgggtggcagcgtgagcgcg	346
Contig40 gene 51	atgatttgcctgctgctgaccttagcaccctgagcgcgattgatgatggcaaacctgacc	347
Contig40 gene 54	atgattattcggattattttatgtataaacgcgtgctgcaacaacgcg	348
Contig40 gene 63	atgaacaagtgcagctgagcagcattctggcctgggtgctgattctgttctgagcctggcgggtggtagcgcg	349
Contig40 gene 70	atgcgcaaaagaaattatagcattctgtgattgagatttgcgattatgagattatgagcagccgcttttagc	350
Contig40 gene 72	atgaacaaaaaacctttaaactgctgctgaccattttattgcgtttgcctgattataacaocigtcttattctgaa	351
Contig40 gene 75	cgataaacctgagc	352
Contig40 gene 105	atgattggtgattctgtgattaccctgctgagcgtgcccgttctgagcctgaccattgattatagc	353
Contig40 gene 119	atgaaacattaacctgaaaaaattacaccttctgtgacctggctggctgattggcctgattagctttaaacagcat	354
Contig40 gene 161	atgaaacaccagaaataatagcgtgcaattgtgtggcggcgcgcaatgagccgcgcgcatgggcccag	355
Contig40 gene 163	atgaaacattctgattaacggcacccggcgcgattggcattggcctggcctggcgtttatgtgcagcagcaccogtgtt	356
Contig40 gene 164	atgattattgtgacccaccttgctgctgattctgattctgattgtgctgttttattggcctgtttccggggcctgaccaa	357
Contig40 gene 165	cagc	358
Contig40 gene 169	atgagcgtgtggcaaacctgattaccaccattataccctggtgaccaccgctttggc	339
Contig40 gene 179	atgaaagcgataaacgcgcgaaatttgcgatttttttagcattgcgattctggcctggcctggcagcaacattgc	340
Contig40 gene 187	ggcg atgattaacggcattatgataaacagaaagtattaccgcgtttggcattatctggtttctggcgcgccgttttag	341
Contig40 gene 203	cccgtttgggtgctgccgattctggccgtg atgttaacaaaaaattgctgctggcgcattagccctgctggcggtgatcttttgcgagcagatgcatgtgagcgcg	342
Contig40 gene 228	atgaaacacacctgaaaaaaccacatttctgctgctgctgctgattctgcgattctgagcattggcagcagat	343
Contig40 gene 231	tagcgcg atgaaacgaaaggcaaatctcgtgctgctgctgattctgattctgagccttagcattattagcgcagcctttgc g atgaaaaaaacctgagcctgaaaaaacattctgattctgagcctgattttctgtttgtgctgagcattggcagcag	344
	ctttgcg	

FIG. 9-2

Contig40 gene 232	atgaaacgcaaccatttatttattattctgctggtgacctggttctgattagcatgagcgtggtgagcgcg	345
Contig40 gene 248	atgaaaaaaaaatggaaatggcagcctataattctctgattgagcgctgctggcattctctgattgacg	346
Contig40 gene 251	atgcccgaattggaacctgtdgaacaaactggcggatccgaataaacatcccgccctgtttgagggtgatt	347
Contig40 gene 252	atgtttaacctggcgattgggtatctggtccctggcctggcatttttggcagccctggcagcctgfgggggccc	348
	ggcgtgaaagtaccg	
Contig40 gene 260	atgtttgcgattgtgagcctgagcgcgtgagcgcgagcgatgatttttagcagcagccctggcggatgatagcgatag	349
	cgat	
Contig40 gene 269	atgaaacgcgcctataaaagtgtgtttctgctggcgattctgaccattattagcattaacgcgattagcgcg	350
Contig40 gene 296	atgctgtttagcgtgatggaacogtggacgctgagcggaccctgc	351
Contig40 gene 297	atgtttattaataatccgcgcgataccctgatttattctgctg	352
Contig40 gene 306	atgaaagcgtgattccggcggggctgggcaaccgctttctgcccggcacaagaagcagcgaagaaatgct	353
	gccggtgatgataaacgcaccattcag	
Contig40 gene 317	atgattaaaaccgatgtgctggtgattgagcggggcggccggcgctttccgcggcacaagaagcagcgaagaaatgct	354
	c	
Contig40 gene 346	atggaagataacctgctgaaataaccgcaactgattctgattagcattttctggtgagcctgctggcgattagcgc	355
	gctgagcgcg	
Contig40 gene 349	atgaaaccgcaacaataattgtgctgctggtgctgctgattgctggtggtggtggtggtggttaccatgggcccggcgtgcgc	356
	g	
Contig40 gene 352	atgaaaaaaagcgtgttaaaattctgattgagcctggcgcctgattctgctggcgtgagcaattgtgagcaagc	357
Contig40 gene 411	atgaaaaaaaaaaacattttctgattgagcatttctgattcggtggtggcgtgagcctggcctgc	358
Contig40 gene 431	atgctgattgctgctgctggcctgagcgcggggcggcggtggattgggatccgcctgaccgataaac	399
Contig40 gene 448	atgagcgaataaacaccgcctgattaccattggcattggcggctttattatt	400
Contig40 gene 466	atgggcataaattttataaattgtgaccaattattctgattgattgagcctggcgttctggtgatttattatag	401
	cgatggccatagcga	
Contig40 gene 483	atggataaaaaaacattatttgcggcgtggcggattctgtgatt	402
Contig40 gene 501	atgaaactgacaataatttttattatagcattattcbgatttctgagcattagcgcgattagcgcg	403
Contig40 gene 553	atgaaaaaaaaatggcatttctggcattgagcctgctggccttctgtgattggcgcgagcagcgcg	404
Contig40 gene 636	atgaaacttaacaaaaaatactgctgatttctgctggcctgtgttattgagcgtggcattgtggcgcgcg	405
Contig40 gene 721	atgaaaccgagcattattttctgaccatttctgagcctgttctgtgattggctatggcagcgcg	406
	atggcattacccttaccgcgattataccggcgcgttatccggcgcgtggcggcaccaccttagcgaaccctgggcaactatct	407
Contig40 gene 730	gagccagtttattccgtat	
Contig40 gene 732	atggatagcaaaaaactgattctggtgacccggcctggctttctggcgttctggcattgtgacattgagcagcgtgagcgcg	408
Contig40 gene 733	atgaaactgacaataaataatttctgctggtgattttattattagcattagcattgagcggcgttattggcggga	409

FIG. 9-4

	catttgaaagcg	
Contig40_gene_1008	atgattctgattattagccgttctctgattagccgtggtggcgattggcggcgagcgcg	436
Contig40_gene_1021	atgaaactgataaaaacagcatattattctgctgattctgattctgagcattggcggcgagcgcg	437
Contig40_gene_1025	atggcggtgattctgatttctgttttagcctggcaccctggcggagcgaacattgtgatt	438
Contig40_gene_1029	atgcaaacccgaaagattatattatgaaaaacgattatctgatttctgattggcggctgctgattagcattgtg	439
Contig40_gene_1036	gagcccgattgcgcg	440
Contig40_gene_1036	atgacaacaacaaatctgtgcccctgctggcattctggcattggcagctggcg	441
Contig40_gene_1074	atgaaaaaacatgaaaagcaaacgttttctctgctgattattattagcattctgattagcaggcgtgag	441
Contig40_gene_1084	cgcg	442
Contig40_gene_1084	atggataaaaaatcttattgagccttattctgctggcattttaccattggcgggtggcgcg	443
Contig40_gene_1088	atgatgaaatgaccacaaaacctgttctgattagcctgattctgctgattctgaccattggcgggtgag	443
Contig40_gene_1089	cgcg	444
Contig40_gene_1089	atggtgattatgacaacaaaactgtttattgtagcctgattattctgaccattctgaccattggcgggtgag	444
Contig40_gene_1093	cgcg	445
Contig40_gene_1093	atgaaatttaacaaacccggcattagcgagattagcattattctgattctgtttctgagcattagcattggcgag	445
Contig40_gene_1097	cgcg	446
Contig40_gene_1097	atggttattctgaaatttgaaatttaaacgacgacctgattttatttagcattctggcgattctgattctgagcattgg	446
Contig40_gene_1098	catggcgagcg	447
Contig40_gene_1098	atgcggcgtattcctccgtgaaagataacttctctgattctggtgacacacatgaaaaaacgattttaaaagcat	447
Contig40_gene_1099	ttttattgctgtgctgctgacctgacctgattggcgggtgagcg	448
Contig40_gene_1104	atgaaacttaaaaaactgctgattagcctgattctgctgtttgtgtgagcgtggcctttagcacccggagcg	448
Contig40_gene_1104	g	449
Contig40_gene_1106	atgaaatttaaaaaagctttgattctgtgctgattattgctgtttaccattggcggcg	449
Contig40_gene_1106	atgaccgtgagcgtgtttattagcgcgcttctgctgttggcaactgctgagcaacgagataacggc	450
Contig40_gene_1158	atgaaagtgtgaaaattgagattattatgctgattctgattctgattagcctggcgggtgagcg	451
Contig40_gene_1176	atgaaacttaaaaaccaaaggcactgattctgatttagcctgctgtttattctgatttggcattggcattggcgag	452
Contig40_gene_1176	cgcg	453
Contig40_gene_1198	atggcaaaatttaaatcttattttattctggtgctggcgtgtttctgatttggcatttggc	453
Contig40_gene_1215	atggtatgcaaaaataattctgattgctgggtggcggctgattgctgattgtggcgggtgagcagctgacgcg	454
Contig40_gene_1238	atgaaactgaaaagcaaatatttgtgttctgctgatttggcattctgattttagcatttagcaccgtgagcg	455
Contig40_gene_1247	atgaaactatagcattattattttttattttttctgattggtgctgctgattggcagcattcaaggctgagcg	456
Contig40_gene_1254	cgcg	457
Contig40_gene_1254	atgaaatttaacagccgctgctggcattctgagcctgctgtttgtgctgaccattctggtgagcagcgtggcg	457

FIG. 9-5

Contig40 gene 1270	atggaaaaaaaaaccaccattattctggtgattctgattgctgctgattcggtgc	458
Contig40 gene 1274	atgaaaaaaaaaagcctgattcttgattagcctgctgctgattaccattatgacattggcagcgtggtggcg	459
Contig40 gene 1296	atgcaagcaccattctgctgagcgcagaccgcgaaagccgagccgcagcctgaccacccgcgcg	460
Contig40 gene 1331	atgctgctgatttcttattggcctggtggaagcattctgaatggcctggtgattgggaagatctggcgattag	461
Contig40 gene 1350	atgcaaaaaaaaaattatctgagctgctgctgctggtgctgctggtggcgatttagcgtgagcgggtggcgccg	462
Contig40 gene 1351	atgagcctgagcattttgtgctgggtgattggcggggttatttaacaaacgattctgctgatttttgtggttct	463
Contig40 gene 1355	gattttttttattagcattggaagcctggcggtggcg	464
Contig40 gene 1364	atgacaacaacacaaaatattatgagcttctgctgctgctgctgattggattagcgtgagcgcgggtgagcgcg	465
Contig40 gene 1367	atgaaaattcagcggcgatattatattctgacctgctgctgctgtttagcctgagcgcggcgagcgcggcg	466
Contig45 gene 8	atggaactgaaagtggatcaagataaactgctggctggctgagtgctgattgctggctggcgggtgaaacgagcgaat	467
Contig45 gene 20	tagccgggaacac	468
Contig45 gene 21	atgaacgcgacgcaaaactgattattcgacattctggtgattctgctgctggcctgctg	469
Contig45 gene 30	atgaaaatataaaaattggcattttggcgcgcccggcgccgagcattgctgctgctggcggaattct	470
Contig45 gene 35	ggcccgaaacgcggtgtgcatctctgaaa	471
Contig45 gene 36	atggcgaaacgaaaggcgcccatctgaaacaccattctgatttattctgattgctgctgatttgctgctg	472
Contig45 gene 60	atggataacacaaaatgaaaggcatttgcctggcgataattgctgctggtggcggtgattggcttatttaa	473
Contig45 gene 64	cgcg	474
Contig45 gene 89	atgaaacctgagaaaaattaccgtggcggtggctggtggtggcgctgagcatttctgctggcgcgaacaca	475
Contig45 gene 91	tgatgtgaccgcg	476
Contig45 gene 93	atggaaaattcgcctataaaaaacctgctgaaaagtgtttaccattttctggtgctgctgattagctgagcgtttgagag	477
Contig45 gene 100	cgcg	478
Contig45 gene 106	atggcaagcagccctgttggataaagtgaaaaccagcctggtgattgctgctgagcgtttttggcctggtg	479
Contig45 gene 116	atgaaaatttaaaacagccatattctgctggtggcctgattgacatatttctgctgctgctgagcattagcgcggcgag	480
Contig45 gene 159	atgcaacagcaaaaattgagctgctgggcattattctgctgagctttgcgatttgagcgagcgcgagcgcg	481
Contig45 gene 159	atggatgatgcacactgctgctgattggctttggcggcgggtggcgaaggcctggcgcgctgctgagcagcgcg	481

FIG. 9-6

	agaaatgattaacgcaaaaatttggcattagcctgaaagtg	
Contig47_gene_7	atgatgcgcaaaacattttggcgtgattttattgtgtttattctgttagcattagcacccgtgagcgcg	482
Contig47_gene_13	atgaacaacagaacgtgttgcctgattctgctgacccaattctctgctgagcgtggtggcgtgagc	483
Contig47_gene_57	atgctgaacaataaataattattctgacattctctgattctgagcattagcagcgcgagcgcgagcgcg	484
Contig47_gene_60	atggcgtgctgctgagcgtggcggcgccatttttttgaagcggcattgattgcgacctgcacc	485
Contig47_gene_62	atgttagcgtgagcctgacaaaactgaaaattggccgcgtttatttgcctgtttattctggtttattagcctg	486
	cagcattaaactgcgtgtttg	
Contig47_gene_98	atggcggcgaatttatacaacgaaaactgaaaactgattctgattctgacc	487
	atggataaaaaaatgattgagcgtggcgttctgctgctgattctggcgtggcgtggtgagcgtgttgatga	488
Contig47_gene_125	aagcaacagc	
Contig47_gene_140	atgaaaaatagcgcattattgtcctgctgctgattctgattcttaccggcggcattggtgtatgctg	489
Contig47_gene_146	atggatagcaaaaaaattctggtgattctggcctgacccgtgctggcattttctggcagcagcgtgagcgcg	490
	atgaaaccgtatgtgattctgattggcagcggcgtgctatcgaccattggcaaaagcaccctggcggcgaaactggcga	491
Contig47_gene_160	cctgaacattaaacatctggtg	
Contig47_gene_197	atgctgattagcgtgctggcgtgattgtgattattattalgggtggcggc	492
Contig47_gene_208	atggcagcccgcaaaaacagaccatttggcgcctgctgatagcattctcagcggcggcctttctg	493
Contig47_gene_253	atggaactgagcaaaagcagataaatactgattgtgggtggcattatttttgcctggcgcgtg	494
Contig47_gene_269	atgaaagtggcgtattctggcggcgtgctatcgaccattggcagcggcattaccacattt	495
Contig47_gene_304	atgsgcgtgattctgattctgtttctggcgtgagcaccgtggcggc	496
Contig47_gene_306	atgcaaaaaaactgaaaaatttctgatactctgctgctgctgatt	497
	tgtgaacatgggcaac	
contig47_gene_309	atgctggtgctggcgtttgcgattatttttctgggtatagcattagcctgggcaacaaccag	498
Contig47_gene_349	atggaactgacgatgaaattattttaaagtggcgtgattaccgcgtggtggcattgctggc	499
Contig47_gene_353	atgaaaaactatttgcataaaagataaagtggtgggtggtagccggcgcgagcggcctggctggcagatt	500
Contig47_gene_356	atgacctttacaaccctgcctaaacattaaagattgattgatttttgggtttaccctgctgctgag	501
Contig47_gene_375	cattctggcgtgagcgcg	502
Contig47_gene_381	atgaaattaaaaatctgtttattctgctgattgctgctgatttgcattattagcgtgagcggctggcgcg	503
Contig47_gene_391	atggataaaaaatgaccgtgctgctggtggcgtgtttgcctgctgctgctggcgcg	504
	atggatcgcaagatattattattctggtgctgattattattagcctgctggcgtggcctggcataacca	505
Contig49_gene_3	tcaggtg	
	atgaccagcaaatattgattctgaccaccgacccggctgctggcggcggatagcgggtgaccattagcgtat	506
Contig49_gene_4	taaa	

FIG. 9-8

NO:117, optimised coding sequence: For consensus SEQ ID		
NO:119, predicted coding sequence	AAA AAA ATT ATT ATT ATT TTA TTA TTA TTA ATT TTA TTA TTA ATT TCA ATT	532
For consensus SEQ ID NO:119, optimised coding sequence	AAA AAA ATT ATT ATT ATT CTG CTG CTG ATT CTG CTG CTG ATT AGC ATT	533

REFERENCIAS CITADAS EN LA DESCRIPCIÓN

Esta lista de referencias citadas por el solicitante es únicamente para la comodidad del lector. No forma parte del documento de la patente europea. A pesar del cuidado tenido en la recopilación de las referencias, no se pueden 5 excluir errores u omisiones y la EPO niega toda responsabilidad en este sentido.

Documentos de patentes citados en la descripción

- 10 • WO 9519763 A [0115]
- NZ 278977 [0115]
- AU 35908178 [0115]
- AU 81100082 W [0115]
- WO 9511995 A, Chee [0127]
- WO 95251116 A, Baldeschweiler [0128]
- WO 8403564 A [0130]

15 Literatura diferente de patentes citada en la descripción

- 20 • DIEFFENBACH, C. W. ; G. S. DVEKSLER. PCR Primer, a Laboratory Manual. Cold Spring Harbor Press, 1995 [0028]
- SAMBROOK, J. et al. Molecular Cloning, A Laboratory Manual. Cold Spring Harbor Press, 1989 [0047] [0071]
- 25 • AUSUBEL, F. M. et al. Current Protocols in Molecular Biology. John Wiley & Sons, 1989 [0047] [0071]
- WAHL, G. M. ; S. L. BERGER. *Methods Enzymol.*, 1987, vol. 152, 399-407 [0063]
- 30 • KIMMEL, A. R. *Methods Enzymol.*, 1987, vol. 152, 507-511 [0063]
- SARKAR, G. *PCR Methods Applic.*, 1993, vol. 2, 318-322 [0065]
- CARUTHERS, M. H. et al. *Nucl. Acids Res. Symp. Ser.*, 1980, 215-223 [0069]
- 35 • HORN, T. et al. *Nucl. Acids Res. Symp. Ser.*, 1980, 225-232 [0069]
- ROBERGE, J. Y. et al. *Science*, 1995, vol. 269, 202-204 [0069]
- 40 • MERRIFIELD J. *J. Am. Chem. Soc.*, 1963, vol. 85, 2149-2154 [0069]
- CREIGTITON, T. *Proteins Structures and Molecular Principles*. WH Freeman and Co, 1983 [0070]
- 45 • VAN HEEKE, G. ; S. M. SCHUSTER. *J. Biol. Chem.*, 1989, vol. 264, 5503-5509 [0074]
- GRANT et al. *Methods Enzymol.*, 1987, vol. 153, 516-544 [0075]
- 50 • SCHARF, D. et al. *Results Probl. Cell Differ.*, 1994, vol. 20, 125-162 [0076]
- FELGNER et al. *Proc. Natl. Acad. Sci., USA*, 1987, vol. 84, 7413-7417 [0078]
- BOTHWELL et al. *Methods for Cloning and Analysis of Eukaryotic Genes*. 1990 [0078]
- 55 • AUSUBEL et al. *Short Protocols in Molecular Biology*. John Wiley and Sons, 1992 [0078]
- FARHOOD. *Annal. NY Acad. Sci.*, 1994, vol. 716, 23-34 [0078]
- 60 • VATTERONI et al. *Mutn. Res*, 1993, vol. 291, 163-169 [0078]
- SABELNIKOV. *Prog. Biophys. Mol. Biol.*, 1994, vol. 62, 119-152 [0078]
- DAVIS et al. *J. Virol.*, 1996, vol. 70 (6), 3781-3787 [0078]
- BRINSTER et al. *J. Gen. Virol.*, 2002, vol. 83 (2), 369-381 [0078]
- MOSS. *Dev. Biol. Stan.*, 1994, vol. 82, 55-63 [0078]
- JOHNSTON et al. *Meth. Cell Biol.*, 1994, vol. 43 (A), 353-365 [0078]
- WATANABE et al. *Mol. Reprod. Dev.*, 1994, vol. 38, 268-274 [0079]
- ROBINSON et al. *Vacc.*, 1993, vol. 11, 957-960 [0079]
- HOFFMAN et al. *Vacc.*, 1994, vol. 12, 1529-1533 [0079]
- XIANG et al. *Virol.*, 1994, vol. 199, 132-140 [0079]
- WEBSTER et al. *Vacc.*, 1994, vol. 12, 1495-1498 [0079]
- DAVIS et al. *Vacc.*, 1994, vol. 12, 1503-1509 [0079]
- DAVIS et al. *Hum. Molec. Gen.*, 1993, vol. 2, 1847-1851 [0079]
- DALEMANS et al. *Ann NY Acad. Sci.*, 1995, vol. 772, 255-256 [0079]
- CONRY et al. *Cancer Res.*, 1995, vol. 55 (7), 1397-1400 [0079]
- NAITO et al. *Mol. Reprod. Dev.*, 1994, vol. 39, 153-161 [0079]
- BURDON et al. *Mol. Reprod. Dev.*, 1992, vol. 33, 436-442 [0079]
- DAVIS et al. *J Virol*, 1996, vol. 70 (6), 3781-3787 [0079]
- BALASURIYA et al. *Vaccine*, 2002, vol. 20 (11 12), 1609-1617 [0079]
- HAMPTON, R. et al. *Serological Methods, a laboratory Manual*. APS Press, 1990 [0080]
- MADDOX, D. E. et al. *J. Exp. Med.*, 1983, vol. 158, 1211-1216 [0080]
- PORATH, J. et al. *Prot. Exp. Purif.*, 1992, vol. 3, 263-281 [0084]
- KROLL, D. J. et al. *DNA Cell Biol.*, 1993, vol. 12, 441-453 [0084]

- KOHLER, G. et al. *Nature*, 1975, vol. 256, 495-497 [0088]
- KOZBOR, D. et al. *J. Immunol. Methods*, 1985, vol. 81, 31-42 [0088]
- 5 • COTE, R. J. et al. *Proc. Natl. Acad. Sci.*, 1983, vol. 80, 2026-2030 [0088]
- COLE, S. P. et al. *Mol. Cell Biol.*, 1984, vol. 62, 109-120 [0088]
- 10 • ORLANDI, R. et al. *Proc. Natl. Acad. Sci.*, 1989, vol. 86, 3833-3837 [0088]
- WINTER, G. et al. *Nature*, 1991, vol. 349, 293-299 [0088]
- 15 • MORRISON, S. L. et al. *Proc. Natl. Acad. Sci.*, 1984, vol. 81, 6851-6855 [0089]
- NEUBERGER, M.S. et al. *Nature*, 1984, vol. 312, 604-608 [0089]
- TAKEDA, S. et al. *Nature*, 1985, vol. 314, 452-454 [0089]
- 20 • TODOROVSKA et al. Design and application of diabodies, triabodies and tetrabodies for cancer targeting. *J. Immunol. Methods*, 01 February 2001, vol. 248 (2), 47-66 [0090]
- 25 • HUSE, W. D. et al. *Science*, 1989, vol. 254, 1275-1281 [0091]
- HANCOCK ; CHAPPLE. *Antimicrob. Agents Chemother.*, 1999, vol. 43, 1317-1323 [0098]
- GANZ ; LEHRER. *Curr. Opin. Hematol.*, 1997, vol. 4, 53-58 [0098]
- 30 • HANCOCK et al. *Adv. Microb. Physiol.*, 1995, vol. 37, 135-175 [0098]
- HORN et al. *Applied Environ. Microbiol.*, 2003, vol. 69, 74-83 [0099]
- 35 • ECKBURG et al. *Infection Immunity*, vol. 71, 591-596 [0099]
- GIJZEN et al. *Applied Environ. Microbiol.*, 1991, vol. 57, 1630-1634 [0099]
- 40 • BONELO et al. *FEMS Microbiol. Lett.*, 1984, vol. 21, 341-345 [0099]
- HUSER et al. *Arch. Microbiol.*, 1982, vol. 132, 1-9 [0099]
- 45 • HILPERT et al. *Zentbl. Bakteriologie. Mikrobiologie. Hyg. 1 Abt Orig. C*, 1981, vol. 2, 21-31 [0099]
- PNA and Oligonucleotide Inhibitors of Human Telomerase. G. GAVORY ; S. BALASUBRAMANIAN. Eureka Bioscience Collection. Landes Bioscience, 2003 [0102]
- 50 • P.E. NIELSEN et al. *Science*, 1991, vol. 254, 1497-1500 [0102]
- M. EGHOLM et al. *Nature*, 1993, vol. 365, 566-568 [0102]
- 55 • V. DEMIDOV et al. *Biochem Pharmacol*, 1994, vol. 48, 1310-1313 [0102]
- NIELSEN, P. E. et al. *Anticancer Drug Des.*, 1993, vol. 8, 53-63 [0102]
- 60 • Remington's Pharmaceutical Sciences. Maack Publishing Co, [0107]
- LU et al. *AAPS J*, 2006, vol. 8, 466-472 [0113]
- DUBIN et al. *Acta Biochemica Polonica*, 2005, vol. 52, 633-638 [0113]
- YEAMAN ; YOUNT. *Pharmacol Rev*, 2003, vol. 55, 27-55 [0113]
- DEGIM et al. *Curr Pharm Des*, 2007, vol. 13, 99-117 [0113]
- SHAI et al. *Curr Prot Pept Sci*, 2006, vol. 7, 479-486 [0113]
- LABY et al. *Can. J. Anim. Sci.*, 1984, vol. 64, 337-8 [0115]
- MELBY, P. C. et al. *J. Immunol. Methods*, 1993, vol. 159, 235-244 [0126]
- DUPLAA, C. et al. *Anal. Biochem.*, 1993, 229-236 [0126]
- LOCKHART, D. J. et al. *Nat. Biotech.*, 1996, vol. 14, 1675-1680 [0127]
- SCHENA, M. et al. *Proc. Natl. Acad. Sci.*, 1996, vol. 93, 10614-10619 [0127]
- SALZBERG et al. *Gene Locator Interpolated Markov Model ER*, 1998 [0135]
- ALTSCHUL et al. *Basic Local Alignment Search Tool*, 1997 [0135]
- PathwayVoyager: pathway mapping using the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. ALTERMANN E ; KLAENHAMMER TR. *BMC Genomics*. 2005, vol. 6, 60-66 [0149]
- ALTERMANN, E. ; T. R. KLAENHAMMER. GAMOLA: a new local solution for sequence annotation and analyzing draft and finished prokaryotic genomes. *Omics*, 2003, vol. 7, 161-169 [0149]
- ALTSCHUL SF ; MADDEN TL ; SCHAFFER AA ; ZHANG J ; ZHANG Z ; MILLER W ; LIPMAN DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, 1997, vol. 25, 3389-3402 [0149]
- BALCHWE ; FOXGE ; MAGRUM LJ ; WOESE CR ; WOLFE RS. Methanogens: reevaluation of a unique biological group. *Microbiological Reviews*, 1979, vol. 43, 260-296 [0149]
- BARESI, L. ; BERTANI, G. Isolation of a bacteriophage for a methanogenic bacterium. In Abstracts of the Annual Meeting of the American Society for Microbiology. *American Society for Microbiology*, 1984, 133 [0149]
- BENDTSEN, J.D. ; NIELSEN, H. ; VON HEIJNE, G. ; BRUNAK, S. Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol.*, 2004, vol. 16 (4), 783-95 [0149]
- BICKLE, T. A. ; D. H. KRUGER. Biology of DNA restriction. *Microbiol. Rev.*, 1993, vol. 57, 434-450 [0149]
- BULT CJ et al. Complete genome sequence of the methanogenic archaeon. *Methanococcus jannaschii*. *Science*, 1996, vol. 273, 1058-1073 [0149]
- Carbohydrate-active enzymes: an integrated database approach. COUTINHO PM ; HENRISSAT B. Recent Advances in Carbohydrate Bioengineering. The Royal Society of Chemistry, 1999, 3-12 [0149]

- DELCHER AL ; HARMON D ; KASIF S ; WHITE O ; SALZBERG SL. Improved microbial gene identification with GLIMMER. *Nucleic Acids Research*, 1999, vol. 27, 4636-4641 [0149]
- 5 • EMANUELSSON, O. ; BRUNAK, S. ; VON HEIJNE, G. ; NIELSEN, H. Locating proteins in the cell using TargetP, SignalP and related tools. *Nat Protoc.*, 2007, vol. 2 (4), 953-71 [0149]
- 10 • FLEISCHMANN et al. Whole-genome random sequencing and assembly of *Haemophilus influenzae*. *Rd Science*, 1995, vol. 269, 496-512 [0149]
- 15 • FRICKE WF ; SEEDORF H ; HENNE A ; KRUEM M ; LIESEGANG H ; HEDDERICH R ; GOTTSCHALK G ; THAUER RK. The genome sequence of *Methanosphaera stadtmanae* reveals why this human intestinal archaeon is restricted to methanol and H₂ for methane formation and ATP synthesis. *Journal of Bacteriology*, 2006, vol. 188, 642-658 [0149]
- 20 • GALAGAN et al. The genome of *Methanosarcina acetivorans* reveals extensive metabolic and physiological diversity. *Genome Res.*, 2002, vol. 12, 532-542 [0149]
- 25 • GODDE JS ; BICKERTON A. The repetitive DNAe called CRISPRs and their associated genes: evidence of horizontal transfer among prokaryotes. *Journal of Molecular Evolution*, 2006, vol. 62, 718-729 [0149]
- 30 • HAFT DH ; SELENGUT J ; MONGODIN EF ; NELSONKE. A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. *PLoS Computational Biology*, 2005, vol. 1, 474-483 [0149]
- 35 • HAMILTON, P. T. ; J. N. REEVE. Sequence divergence of an archaeobacterial gene cloned from a mesophilic and a thermophilic methanogen. *J. Mol. Evol.*, 1985, vol. 22, 351-360 [0149]
- 40 • JANSEN R ; EMBDEN JD ; GAASTRA W ; SCHOULS LM. Identification of genes that are associated with DNA repeats in prokaryotes. *Molecular Microbiology*, 2002, vol. 43, 1565-1575 [0149]
- 45 • JANSEN R ; VAN EMBDEN JD ; GAASTRA W ; SCHOULS LM. Identification of a novel family of sequence repeats among prokaryotes. *OMICS: A journal of integrative biology*, 2002, vol. 6, 23-33 [0149]
- 50 • JENSEN, L. J. ; FRIIS, C. ; USSERY, D. W. Three views of microbial genomes. *Res. Microbiol.*, 1999, vol. 150, 773-777 [0149]
- 55 • JOBLIN KN ; NAYLOR GE ; WILLIAMS AG. Effect of *Methanobrevibacter smithii* on xylanolytic activity of anaerobic ruminal fungi. *Applied and Environmental Microbiology*, 1990, vol. 56, 2287-2295 [0149]
- 60 • KANEHISA M ; GOTO S ; KAWASHIMA S ; OKUNO Y ; HATTORI M. The KEGG resource for deciphering the genome. *Nucleic Acids Research*, 2004, vol. 32, 277-D280 [0149]
- KIENER, A. ; KONIG, H. ; WINTER, J. ; LEISINGER, T. Purification and use of *Methanobacterium wolfei* pseudomurein endopeptidase for lysis of *Methanobacterium thermoautotrophicum*. *J. Bacteriol.*, 1987, vol. 169, 1010-1016 [0149]
- Isolation and characterisation of a bacteriophage of *Methanobrevibacter smithii*. KNOX, M. R. ; HARRIS, J. E. In Abstracts of the XIV International Congress on Microbiology. International Union of Microbiological Societies, 1986 [0149]
- KURTZ S ; SCHLEIERMACHER C. REPuter: fast computation of maximal repeats in complete genomes. *Bioinformatics*, 1999, vol. 15, 426-427 [0149]
- LARKIN, M.A. ; BLACKSHIELDS, G. ; BROWN, N.P. ; CHENNA, R. ; MCGETTIGAN, P.A. ; MCWILLIAM, H. ; VALENTIN, F. ; WALLACE, I.M. ; WILM, A. ; LOPEZ, R. *Bioinformatics*, vol. 23 (21), 2947-8 [0149]
- LOWE TM ; EDDY SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Research*, 1997, vol. 25, 955-964 [0149]
- LOENEN, W. ; N. MURRAY. Modification enhancement by restriction alleviation protein (Ra1) of bacteriophage lambda. *J. Mol. Biol.*, 1986, vol. 190, 11-22 [0149]
- LUCCHINI, S. ; F. DESIERE ; H. BRUSSOW. Comparative genomics of *Streptococcus thermophilus* phage species supports a modular evolution theory. *J. Virol.*, 1999, vol. 73, 8647-8656 [0149]
- LUO, Y. N. ; PFISTER, P. ; LEISINGER, T. ; WASSERFALLEN, A. Pseudomurein endoisopeptidases PeiW and PeiP, two moderately related members of a novel family of proteases produced in *Methanothermobacter* strains. *FEMS Microbiol. Lett.*, 2002, vol. 208, 47-51 [0149]
- MAKAROVA, K. S. ; ARAVIND, L. ; KOONIN, E. V. A superfamily of archaeal, bacterial, and eukaryotic proteins homologous to animal transglutaminases. *Protein Sci.*, 1999, vol. 8, 1714-1719 [0149]
- MAKAROVA KS ; GRISHIN NVSHABALINA, SA ; WOLF YI ; KOONIN EV. A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biology Direct*, 2006, vol. 1, 7-32 [0149]
- NIELSEN, H. ; ENGELBRECHT, J. ; BRUNAK S. ; VON HEIJNE, G. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Prot. Eng.*, 1997, vol. 10, 1-6 [0149]
- ONO, M. ; WADA, Y. ; WU, Y. ; NEMORI, R. ; JINBO, Y. ; WANG, H. ; LO, K. M. ; YAMAGUCHI, N. ; BRUNKHORST, B. ; OTOMO, H. et al. FP-21399 blocks HIV envelope protein mediated membrane fusion and concentrates in lymph nodes. *Nat. Biotechnol.*, 1997, vol. 15, 343-348 [0149]

- **PEREZ-BERCOFF, A. ; KOCH, J. ; BÜRGLIN, T.R.** LogoBar: bar graph visualization of protein logos with gaps. *Bioinformatics*, 2006, vol. 22 (1), 112-4 [0149]
- 5 • **RAWLINGS, N. D. ; MORTON, F. R. ; BARRETT, A. J.** MEROPS: the peptidase database. *Nucleic Acids Res.*, 2006, vol. 34, 270-D272 [0149]
- **REEVE JN ; NOLLING J ; MORGAN RM ; SMITH DR.** Methanogenesis: genes, genomes and who's on first?. *Journal of Bacteriology*, 1997, vol. 179, 5975-5986 [0149]
- 10 • **SAMUEL BS ; HANSEN EE ; MANCHESTER JK ; COUTINHO PM ; HENRISSAT B ; FULTON R ; LATREILLE P ; KIM K ; WILSON RK ; GORDON JI.** Genomic adaptations of *Methanobrevibacter smithii* to the human gut. *Proceedings of the National Academy of Sciences USA*, 2007, vol. 104, 10643-10648 [0149]
- 15 • **SALZBERG et al.** Microbial gene identification using interpolated Markov models. *Nucleic Acids Res.*, 1998, vol. 26, 544-8 [0149]
- 20 • **SMITH et al.** Complete genome sequence of *Methanobacterium thermoautotrophicum deltaH*: functional analysis and comparative genomics. *J. Bacteriol.*, 1997, vol. 179, 7135-7155 [0149]
- **SMITH PH ; HUNGATE RE.** Isolation and characterization of *Methanobacterium ruminantium* n. sp. *Journal of Bacteriology*, 1958, vol. 75, 713-718 [0149]
- **STADEN R ; BEAL KF ; BONFIELD JK.** The Staden Package. *Methods in Molecular Biology: Bioinformatics Methods and Protocols*, 1998, vol. 132, 115-130 [0149]
- **TATUSOV RL ; NATALE DA ; GARKAVTSEV IV ; TATUSOVA TA ; SHANKAVARAM UT ; RAO BS ; KIRYUTIN B ; GALPERIN MY ; FEDOROVA ND ; KOONIN EV.** The COG database: new developments in phylogenetic classification of proteins from complete genomes. *Nucleic Acids Research*, 2001, vol. 29, 22-28 [0149]
- **VON HEIJNE, G.** Signal sequences: The limits of variation. *J. Mol. Biol.*, 1985, vol. 184, 99-105 [0149]
- 25
- 30
- 35
- 40
- 45
- 50