

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 405 405**

51 Int. Cl.:

**C07K 14/195** (2006.01)

**C07K 14/205** (2006.01)

**G01N 33/68** (2006.01)

**G01N 33/50** (2006.01)

12

## TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **16.12.2002 E 02805613 (3)**

97 Fecha y número de publicación de la concesión europea: **30.01.2013 EP 1581119**

54 Título: **Composiciones y procedimientos para la terapia y el diagnóstico de enfermedad inflamatoria del intestino**

30 Prioridad:

**17.12.2001 US 341830 P**

**16.07.2002 US 396242 P**

**15.11.2002 US 426835 P**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**31.05.2013**

73 Titular/es:

**CORIXA CORPORATION (100.0%)**  
**2711 Centerville Road, Suite 400**  
**Wilmington, DE 19808, US**

72 Inventor/es:

**HERSHBERG, ROBERT M.;**  
**HOSKEN, NANCY ANN;**  
**LODES, MICHAEL J. y**  
**MOHAMATH, RAODOH**

74 Agente/Representante:

**CURELL AGUILÁ, Mireia**

**ES 2 405 405 T3**

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

## DESCRIPCIÓN

Composiciones y procedimientos para la terapia y el diagnóstico de enfermedad inflamatoria del intestino.

5 **Antecedentes de la invención****Campo de la invención**

10 La presente invención se refiere de manera general a la terapia y al diagnóstico de la enfermedad de Crohn y la colitis ulcerosa (denominadas colectivamente enfermedad inflamatoria del intestino, o EII). La invención se refiere más particularmente a polipéptidos que comprenden al menos una parte de una proteína que se reconoce, y frente a la cual individuos con EII montan una respuesta inmunitaria aberrante, y a polinucleótidos que codifican para tales polipéptidos. Tales polipéptidos y polinucleótidos son útiles en composiciones farmacéuticas, por ejemplo, vacunas, y otras composiciones para el diagnóstico y el tratamiento de EII.

15

**Descripción de la técnica relacionada**

20 La enfermedad de Crohn y la colitis ulcerosa (denominadas colectivamente enfermedad inflamatoria del intestino, o EII) son enfermedades inflamatorias crónicas del tubo gastrointestinal. Aunque las características clínicas varían algo entre estos dos trastornos, ambos se caracterizan por dolor abdominal, diarrea (con frecuencia con hemorragia), un grupo variable de manifestaciones "extraintestinales" (tales como artritis, uveítis, cambios en la piel, etc.) y la acumulación de células inflamatorias dentro del intestino delgado y el colon (observada en muestras quirúrgicas o de biopsia patológica).

25 La EII afecta tanto a niños como a adultos y presenta una distribución por edades bimodal (un máximo alrededor de 20 y otro alrededor de 40). La EII es una enfermedad crónica de por vida, y con frecuencia se agrupa con otros trastornos denominados "autoinmunitarios" (por ejemplo artritis reumatoide, diabetes mellitus tipo I, esclerosis múltiple, etc.). La EII se encuentra casi exclusivamente en los países industrializados. Los datos más recientes de la Clínica Mayo sugieren una incidencia global superior a 1 de cada 100.000 personas en los Estados Unidos, con datos de prevalencia en algunos estudios superiores a 1 de cada 1000. Hay una clara tendencia hacia el aumento de la incidencia de EII en los EE.UU. y en Europa, particularmente la enfermedad de Crohn. La base para este aumento no está clara en la actualidad. Como tal, la EII representa la 2ª enfermedad autoinmunitaria más común en los Estados Unidos (tras la artritis reumatoide).

35 El tratamiento de EII es diverso. La terapia de primera línea incluye normalmente derivados de salicilato (por ejemplo 5-ASA) administrados por vía oral o por vía rectal. Las tasas de respuesta en enfermedad de Crohn sin complicaciones son de aproximadamente el 40% (en comparación con el 20% para placebo). Los corticosteroides siguen siendo la piedra angular en el tratamiento de pacientes con enfermedad más "resistente", a pesar de los efectos secundarios desfavorables. Las opciones de tratamiento más recientes incluyen anti-metabolitos (por ejemplo metotrexato, 6-mercaptopurina) e inmunomoduladores (por ejemplo Remicade: un anticuerpo humano quimérico dirigido al receptor de TNF $\alpha$ ).

45 A pesar de la considerable investigación sobre terapias para estos trastornos, la EII sigue siendo difícil de diagnosticar y de tratar eficazmente. Además, no hay pruebas de laboratorio claras para diagnosticar la EII, ni tampoco hay pruebas de laboratorio adecuadas que sirvan como "marcador sustituto" que sean útiles de manera uniforme para realizar un seguimiento del transcurso de la enfermedad en pacientes. Por consiguiente, existe una necesidad en la materia de procedimientos mejorados para detectar y tratar tales enfermedades inflamatorias del intestino. La presente invención cumple estas necesidades y proporciona además otras ventajas relacionadas.

50 **Breve descripción de las figuras**

La figura 1 muestra un esquema de clones de flagelina con porcentaje de similitud con flagelina B relacionada de la bacteria ruminal anaerobia, *Butyrivibrio fibrisolvens*.

55 La figura 2a muestra análisis de inmunotransferencia de tipo Western de respuesta de anticuerpos en suero contra flagelinas recombinantes cBir-1 y Fla<sup>x</sup> y fragmentos.

La figura 2b muestra la titulación de anticuerpo anti-flagelina en suero contra flagelinas recombinantes cBir-1 y Fla<sup>x</sup>.

60 La figura 2c muestra la correlación de puntuación de colitis con anticuerpo anti-Fla<sup>x</sup> en suero.

Las figuras 3a y b muestran la liberación de citocinas por donantes estimulados con flagelina.

**Breve resumen de la invención**

65 En un aspecto, la presente invención proporciona un polinucleótido aislado que comprende una secuencia

seleccionada de entre el grupo que consiste en:

(a) la secuencia proporcionada en SEC ID N°: 75;

5 (b) el complemento de la secuencia proporcionada en SEC ID N°: 75;

(c) secuencias que presentan al menos el 75% o el 90% de identidad con la secuencia de SEC ID N°: 75.

10 La presente invención, en otro aspecto, proporciona un polipéptido aislado que comprende una secuencia de aminoácidos que se codifica por una secuencia de polinucleótido descrita anteriormente.

15 La presente invención proporciona además un polipéptido aislado que comprende una secuencia de aminoácidos seleccionada de entre el grupo que consiste en a) la secuencia mencionada en SEC ID N°: 79; y b) secuencias que presentan al menos el 70%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% o el 99% de identidad con la secuencia expuesta en SEC ID N°: 79.

20 La invención también proporciona un fragmento de polipéptido aislado que comprende al menos 25, 50, 100 o más aminoácidos contiguos de la secuencia de SEC ID N° 79, siendo el fragmento de polipéptido inmunológicamente reactivo con un anticuerpo que se une al polipéptido de longitud completa de SEC ID N° 79.

En determinadas realizaciones preferidas, los polipéptidos y/o polinucleótidos de la presente invención son inmunogénicos, es decir, pueden provocar una respuesta inmunitaria, particularmente una respuesta inmunitaria humoral y/o celular, tal como se describe adicionalmente en la presente memoria.

25 En la presente memoria se dan a conocer fragmentos, variantes y/o derivados de las secuencias de polipéptido y/o de polinucleótido dadas a conocer, presentando los fragmentos, variantes y/o derivados preferentemente un nivel de actividad inmunogénica de al menos aproximadamente el 50%, preferentemente al menos aproximadamente el 70% y más preferentemente al menos aproximadamente el 90% del nivel de actividad inmunogénica de una secuencia de polipéptido expuesta en SEC ID N°: 79 o una secuencia de polipéptido codificada por una secuencia de polinucleótido expuesta en SEC ID N°: 75.

30 La presente invención proporciona además polinucleótidos que codifican para un polipéptido descrito anteriormente, vectores de expresión que comprenden tales polinucleótidos y células huésped transformadas o transfectadas con tales vectores de expresión.

35 La invención proporciona además un procedimiento de preparación de un polipéptido de la invención que comprende cultivar una célula huésped transformada o transfectada con un vector de expresión que comprende un polinucleótido de la invención funcionalmente unido a una secuencia de control de la expresión en condiciones mediante las cuales se expresa el polinucleótido y se produce el polipéptido; y recuperar el polipéptido.

40 Otro aspecto de la presente invención proporciona anticuerpos aislados, o fragmento de unión a antígeno de los mismos, que se unen específicamente al polipéptido de SEC ID N°: 79. En una realización de la invención, el anticuerpo puede ser un anticuerpo monoclonal. En una realización adicional el anticuerpo es un anticuerpo humano o un anticuerpo que se ha humanizado. En realizaciones aún adicionales, los anticuerpos de la presente invención se unen a proteínas de flagelina y en una realización los anticuerpos son anticuerpos neutralizantes contra proteínas de flagelina. En una realización adicional, dichos anticuerpos bloquean la interacción entre una proteína de flagelina y un receptor tipo Toll. En una realización particular, el receptor tipo Toll es TLR5.

45 La presente invención proporciona además, en otros aspectos, proteínas de fusión que comprenden al menos un polipéptido tal como se describió anteriormente, así como polinucleótidos que codifican para tales proteínas de fusión, normalmente en forma de composiciones farmacéuticas, por ejemplo, composiciones de vacuna, que comprenden un portador fisiológicamente aceptable y/o un inmunoestimulante. Las proteínas de fusión pueden comprender múltiples polipéptidos inmunogénicos o partes/variantes de los mismos, tal como se describe en la presente memoria, y pueden comprender además uno o más segmentos de polipéptido para facilitar la expresión, purificación y/o inmunogenicidad del/de los polipéptido(s).

50 En la presente memoria se dan a conocer oligonucleótidos que se hibridan con los polinucleótidos de la presente invención. Los oligonucleótidos pueden hibridarse con los polinucleótidos de la presente invención en condiciones de alta rigurosidad. Los oligonucleótidos pueden hibridarse con polinucleótidos que codifican para proteínas de flagelina.

60 La presente invención proporciona además, en un aspecto, procedimientos de estimulación y/o expansión de células T específicas para una proteína bacteriana entérica, que comprende poner en contacto células T con al menos un componente incluyendo, pero sin limitarse a, polipéptidos o polinucleótidos de la presente invención, células presentadoras de antígeno que expresan un polinucleótido de la presente invención en condiciones y durante un tiempo suficientes para permitir la estimulación y/o la expansión de células T. En una realización de la invención las

células T son células T CD4+. En una realización adicional, las células T CD4+ median una disminución en la inflamación en el colon. En otra realización las células T son específicas para un polipéptido de flagelina.

5 También se describen poblaciones de células T producidas según los procedimientos descritos en la presente memoria. Dichas células T pueden producir citocinas que pueden incluir, pero no se limitan a, interleucina 10 (IL-10), interferón- $\gamma$  (IFN- $\gamma$ ), interleucina 4 (IL-4), interleucina 12 (IL-12), factor de crecimiento transformante beta (TGF $\beta$ ) o interleucina 18 (IL-18). Preferentemente las células T producen IL-10 y/o TGF $\beta$ .

10 Dentro de otros aspectos, la presente invención proporciona composiciones farmacéuticas que comprenden un polipéptido o polinucleótido tal como se describió anteriormente y un portador fisiológicamente aceptable.

15 Dentro de un aspecto relacionado de la presente invención, las composiciones farmacéuticas, por ejemplo, composiciones de vacuna, se proporcionan para aplicaciones profilácticas o terapéuticas. Tales composiciones comprenden generalmente un polipéptido inmunogénico o polinucleótido de la invención y un inmunostimulante, tal como un adyuvante.

20 La presente invención proporciona además composiciones farmacéuticas que comprenden: (a) un anticuerpo o fragmento de unión a antígeno del mismo que se une específicamente a un polipéptido de la presente invención, o un fragmento del mismo; y (b) un portador fisiológicamente aceptable.

25 La presente invención proporciona además, en otros aspectos, proteínas de fusión que comprenden al menos un polipéptido tal como se describió anteriormente, normalmente en forma de composiciones farmacéuticas, por ejemplo, composiciones de vacuna, que comprenden un portador fisiológicamente aceptable y/o un inmunostimulante. Las proteínas de fusión pueden comprender múltiples polipéptidos inmunogénicos o partes/variantes de los mismos, tal como se describe en la presente memoria, y pueden comprender además uno o más segmentos de polipéptido para facilitar la expresión, purificación y/o inmunogenicidad del/de los polipéptido(s).

30 Dentro de aspectos adicionales, la presente invención proporciona composiciones farmacéuticas que comprenden: (a) células T específicas para un polipéptido tal como se describió anteriormente y (b) un portador o excipiente farmacéuticamente aceptable. Las células T ilustrativas incluyen células T que expresan una variedad de citocinas incluyendo interleucina 10 (IL-10), interferón- $\gamma$  (IFN- $\gamma$ ), interleucina 4 (IL-4), interleucina 12 (IL-12), factor de crecimiento transformante beta (TGF $\beta$ ) o interleucina 18 (IL-18). En realizaciones preferidas, las células T producen IL-10 y/o TGF $\beta$ .

35 Dentro de aspectos adicionales, la presente invención proporciona composiciones farmacéuticas que comprenden: (a) una célula presentadora de antígeno que expresa un polipéptido tal como se describió anteriormente y (b) un portador o excipiente farmacéuticamente aceptable. Las células presentadoras de antígeno ilustrativas incluyen células dendríticas, macrófagos, monocitos, fibroblastos y células B.

40 Dentro de aspectos relacionados, se proporcionan composiciones farmacéuticas que comprenden: (a) células T específicas para un polipéptido tal como se describió anteriormente o una célula presentadora de antígeno que expresa un polipéptido tal como se describió anteriormente y (b) un inmunostimulante. Los inmunostimulantes ilustrativos incluyen adyuvantes tales como adyuvante incompleto de Freund; adyuvante completo de Freund; adyuvante 65 de Merck; AS-1, AS-2; gel de hidróxido de aluminio; fosfato de aluminio; una sal de calcio, hierro o zinc; una suspensión insoluble de tirosina acilada, azúcares acilados; polisacáridos derivados catiónica o aniónicamente; polifosfacenos; microesferas biodegradables; monofosforil lípido A, QS21, 4-fosfatos de aminoalquilglucosaminida o Quil A.

50 También se da a conocer un procedimiento de estimulación de una respuesta inmunitaria en un mamífero, que comprende administrar al mamífero las composiciones descritas anteriormente. La respuesta inmunitaria puede comprender células T que producen una citocina incluyendo, pero sin limitarse a, interleucina 10 (IL-10), interferón- $\gamma$  (IFN- $\gamma$ ), interleucina 4 (IL-4), interleucina 12 (IL-12), factor de crecimiento transformante beta (TGF $\beta$ ) o interleucina 18 (IL-18). Las citocinas ilustrativas particulares comprenden IL-10 y/o TGF $\beta$ .

55 También se da a conocer un procedimiento de disminución de inflamación gastrointestinal asociada con enfermedad inflamatoria del intestino en un mamífero, que comprende administrar a dicho mamífero las composiciones de la presente invención.

60 Otro aspecto de la presente invención proporciona un procedimiento de detección de la presencia de enfermedad inflamatoria del intestino en un mamífero que comprende poner en contacto una muestra biológica del mamífero, en el que dicha muestra biológica comprende anticuerpos, con los polipéptidos descritos anteriormente, detectar en la muestra una cantidad de anticuerpo que se une al polipéptido; y comparar la cantidad de anticuerpo unido con un valor de corte predeterminado y a partir de ello determinar la presencia de enfermedad inflamatoria del intestino en el mamífero. Las muestras biológicas ilustrativas incluyen sueros, heces, tejido u otro material obtenido mediante colonoscopia, ileoscopia, esofagogastroduodenoscopia (EGP) o cirugía. En una realización particular el polipéptido comprende un polipéptido de flagelina.

65

También se da a conocer un procedimiento de detección de la presencia de enfermedad inflamatoria del intestino en un mamífero que comprende poner en contacto una muestra biológica del mamífero, en el que dicha muestra biológica comprende polinucleótidos, con al menos un oligonucleótido que es al menos parcialmente complementario a un polinucleótido descrito anteriormente, detectar en la muestra una cantidad de un polinucleótido que se hibrida con dicho oligonucleótido; y comparar la cantidad de dicho polinucleótido que se hibrida con dicho oligonucleótido con un valor de corte predeterminado, y a partir de ello determinar la presencia de enfermedad inflamatoria del intestino en un mamífero. El oligonucleótido puede hibridarse en condiciones de rigurosidad moderada. En particular, puede utilizarse la reacción en cadena de la polimerasa para determinar la cantidad de polinucleótido que se hibrida con dicho oligonucleótido. Puede utilizarse un ensayo de hibridación para determinar la cantidad de polinucleótido que se hibrida con dicho oligonucleótido. Las muestras biológicas que comprenden polinucleótidos ilustrativas incluyen sueros, heces, tejido u otro material obtenido mediante colonoscopia o biopsia de colon, ileoscopia, esofagogastroduodenoscopia (EGP) o cirugía. En particular, el polinucleótido puede codificar para una proteína de flagelina.

Otro aspecto de la presente invención proporciona un procedimiento *ex vivo* de estimulación y/o expansión de células B que producen anticuerpos específicos para una proteína bacteriana entérica, que comprende poner en contacto células B con los polipéptidos o polinucleótidos mencionados anteriormente en condiciones y durante un tiempo suficientes para permitir la estimulación y/o la expansión de células B. En una realización las células B producen anticuerpos que se unen a una proteína de flagelina. En otra realización dichos anticuerpos son anticuerpos neutralizantes contra una proteína de flagelina. En otra realización, los anticuerpos bloquean la interacción entre una proteína de flagelina y un receptor tipo Toll. En una realización particular, el receptor tipo Toll es TLR5.

También se dan a conocer poblaciones de células B generadas tal como se describió anteriormente.

Dentro de aspectos adicionales, la presente invención proporciona un procedimiento de identificación de antígenos bacterianos asociados con enfermedad inflamatoria del intestino en un mamífero, que comprende poner en contacto una muestra biológica que comprende células T con los polinucleótidos o polipéptidos descritos anteriormente, o células presentadoras de antígeno que expresan un polinucleótido de la invención, en condiciones y durante un tiempo suficientes para permitir la estimulación y/o la expansión de células T, y además, detectar en la muestra la magnitud de dicha estimulación y/o expansión de células T; y comparar la magnitud de dicha estimulación y/o expansión con un valor de corte predeterminado, y a partir de ello identificar antígenos bacterianos asociados con enfermedad inflamatoria del intestino en el mamífero. En una realización, el mamífero es un ser humano. En una realización adicional el mamífero es un ratón. Cepas de ratón ilustrativas son C3H/HeJ Bir, BALB/c B6 IL-10<sup>-/-</sup>, B10 IL-10<sup>-/-</sup>, MDR1a<sup>-/-</sup>, TCR $\alpha$ <sup>-/-</sup>, IL-2<sup>-/-</sup>, IL-2R<sup>-/-</sup>, ratones con colitis inducida por DSS (dextrano-sulfato de sodio), G $\alpha_{ai}$ <sup>-/-</sup> y ratones transgénicos CD45 RB. En una realización particular, la cepa de ratón es C3H/HeJ Bir. En otra realización, el mamífero es una rata. En una realización particular, la rata es una rata transgénica HLA-B27.

En determinados otros aspectos, la presente invención proporciona procedimientos de monitorización de la progresión de enfermedad inflamatoria del intestino en un mamífero, que comprenden las etapas de: (a) poner en contacto una muestra biológica obtenida del mamífero, en los que dicha muestra biológica comprende anticuerpos, con un polipéptido de la invención; (b) detectar en la muestra una cantidad de anticuerpo que se une al polipéptido; y (c) repetir las etapas (a) y (b), utilizando una muestra biológica obtenida del mamífero en un punto temporal posterior; y (d) comparar la cantidad de anticuerpo unido en la etapa (b) con la cantidad de anticuerpo unido en la etapa (c) y a partir de ello monitorizar la progresión de enfermedad inflamatoria del intestino en el mamífero.

También se dan a conocer procedimientos de identificación de antígenos bacterianos asociados con enfermedad inflamatoria del intestino en un primer mamífero, que comprenden las etapas de: (a) obtener una muestra biológica de dicho primer mamífero en los que dicha muestra biológica comprende ADN de bacterias cecales; (b) construir una biblioteca de expresión con dicho ADN; (c) examinar dicha biblioteca de expresión con sueros procedentes de dicho primer mamífero o de un segundo mamífero con enfermedad inflamatoria del intestino; identificando así antígenos bacterianos asociados con enfermedad inflamatoria del intestino. Ambos de dichos mamíferos primero y segundo pueden ser ratones. Dicho primer mamífero puede ser un ratón C3H/HeJ Bir y dicho segundo mamífero puede ser una cepa de ratón diferente incluyendo BALB/c IL-10<sup>-/-</sup>, B6 IL-10<sup>-/-</sup>, MDR1a<sup>-/-</sup> o ratones transgénicos CD45 RB. Alternativamente, dicho primer mamífero puede ser un ratón y dicho segundo mamífero puede ser un ser humano.

También se dan a conocer procedimientos de monitorización de la progresión de enfermedad inflamatoria del intestino en un mamífero, que comprenden las etapas de (a) obtener una muestra biológica de dicho mamífero, en los que dicha muestra biológica comprende polinucleótidos; (b) poner en contacto dicha muestra con al menos un oligonucleótido que es al menos parcialmente complementario a un polinucleótido descrito en la presente memoria; (c) detectar en la muestra una cantidad de un polinucleótido que se hibrida con dicho oligonucleótido; (d) repetir las etapas (a), (b) y (c) utilizando una muestra biológica obtenida de dicho mamífero en un punto temporal posterior; y (e) comparar la cantidad de dicho polinucleótido que se hibrida con dicho oligonucleótido en la etapa (c) con la cantidad de dicho polinucleótido que se hibrida con dicho oligonucleótido en la etapa (d); y a partir de ello monitorizar

la progresión de enfermedad inflamatoria del intestino en un mamífero. El oligonucleótido puede hibridarse en condiciones de rigurosidad moderada. En particular, puede utilizarse la reacción en cadena de la polimerasa para determinar la cantidad de polinucleótido que se hibrida con dicho oligonucleótido. Alternativamente, la cantidad de polinucleótido que se hibrida con dicho oligonucleótido puede determinarse utilizando un ensayo de hibridación. Muestras biológicas ilustrativas son sueros, heces, tejido u otro material obtenido mediante colonoscopia o biopsia de colon, ileoscopia, esofagogastroduodenoscopia (EGP) o cirugía. El polinucleótido puede comprender un polinucleótido que codifica para una proteína de flagelina.

Otros aspectos de la presente invención proporcionan kits de diagnóstico. Un kit de diagnóstico puede comprender al menos un anticuerpo de la invención y un reactivo de detección, en el que el reactivo de detección comprende un grupo indicador.

En determinados aspectos, la presente invención proporciona un kit de diagnóstico que comprende al menos uno o más polipéptidos de la invención; y un reactivo de detección que comprende un grupo indicador. En una realización relacionada, dicho al menos uno o más polipéptidos se inmoviliza sobre un soporte sólido. Los reactivos de detección ilustrativos comprenden un anticuerpo anti-inmunoglobulina, proteína G, proteína A o una lectina. Los grupos indicadores ilustrativos comprenden grupos radiactivos, grupos fluorescentes, grupos luminiscentes, enzimas, biotina o tintes.

**Breve descripción de los identificadores de secuencia**

SEC ID N°: 1 es la secuencia de ADNc determinada para 76779.

SEC ID N°: 2 es la secuencia de ADNc determinada para 76780.

SEC ID N°: 3 es la secuencia de ADNc determinada para 76959.

SEC ID N°: 4 es la secuencia de ADNc determinada para 76960.

SEC ID N°: 5 es la secuencia de ADNc determinada para 76961.

SEC ID N°: 6 es la secuencia de ADNc determinada para 76781.

SEC ID N°: 7 es la secuencia de ADNc determinada para 76962.

SEC ID N°: 8 es la secuencia de ADNc determinada para 76782.

SEC ID N°: 9 es la secuencia de ADNc determinada para 76963.

SEC ID N°: 10 es la secuencia de ADNc determinada para 76964.

SEC ID N°: 11 es la secuencia de ADNc determinada para 77529.

SEC ID N°: 12 es la secuencia de ADNc determinada para 76965.

SEC ID N°: 13 es la secuencia de ADNc determinada para 76966.

SEC ID N°: 14 es la secuencia de ADNc determinada para 76967.

SEC ID N°: 15 es la secuencia de ADNc determinada para 76968.

SEC ID N°: 16 es la secuencia de ADNc determinada para 77530.

SEC ID N°: 17 es la secuencia de ADNc determinada para 76969.

SEC ID N°: 18 es la secuencia de ADNc determinada para 76970.

SEC ID N°: 19 es la secuencia de ADNc determinada para 76971.

SEC ID N°: 20 es la secuencia de ADNc determinada para 77073.

SEC ID N°: 21 es la secuencia de ADNc determinada para 76972.

SEC ID N°: 22 es la secuencia de ADNc determinada para 76973.

SEC ID N°: 23 es la secuencia de ADNc determinada para 76974.

- SEC ID Nº: 24 es la secuencia de ADNc determinada para 77074.
- 5 SEC ID Nº: 25 es la secuencia de ADNc determinada para 77531.
- SEC ID Nº: 26 es la secuencia de ADNc determinada para 76975.
- SEC ID Nº: 27 es la secuencia de ADNc determinada para 77075.
- 10 SEC ID Nº: 28 es la secuencia de ADNc determinada para 76976.
- SEC ID Nº: 29 es la secuencia de ADNc determinada para 76977.
- 15 SEC ID Nº: 30 es la secuencia de ADNc determinada para 77532.
- SEC ID Nº: 31 es la secuencia de ADNc determinada para 77533.
- SEC ID Nº: 32 es la secuencia de ADNc determinada para 77534.
- 20 SEC ID Nº: 33 es la secuencia de ADNc determinada para 77535.
- SEC ID Nº: 34 es la secuencia de ADNc determinada para 77076.
- SEC ID Nº: 35 es la secuencia de ADNc determinada para 77536.
- 25 SEC ID Nº: 36 es la secuencia de ADNc determinada para 77538.
- SEC ID Nº: 37 es la secuencia de ADNc determinada para 77539.
- 30 SEC ID Nº: 38 es la secuencia de aminoácidos codificada por 76779.
- SEC ID Nº: 39 es la secuencia de aminoácidos codificada por 76780.
- SEC ID Nº: 40 es la secuencia de aminoácidos codificada por 76959.
- 35 SEC ID Nº: 41 es la secuencia de aminoácidos codificada por 76959.
- SEC ID Nº: 42 es la secuencia de aminoácidos codificada por 76781.
- 40 SEC ID Nº: 43 es la secuencia de aminoácidos codificada por 76782.
- SEC ID Nº: 44 es la secuencia de aminoácidos codificada por 76967.
- 45 SEC ID Nº: 45 es la secuencia de aminoácidos codificada por 76969.
- SEC ID Nº: 46 es la secuencia de aminoácidos codificada por 76972.
- SEC ID Nº: 47 es la secuencia de aminoácidos codificada por 76974.
- 50 SEC ID Nº: 48 es la secuencia de aminoácidos codificada por 76975.
- SEC ID Nº: 49 es la secuencia de aminoácidos codificada por 76977.
- 55 SEC ID Nº: 50 es la secuencia de aminoácidos codificada por 77076.
- SEC ID Nº: 51 es la secuencia de ADNc determinada para 73261.
- SEC ID Nº: 52 es la secuencia de ADNc determinada para 73262.
- 60 SEC ID Nº: 53 es la secuencia de ADNc determinada para 73263.
- SEC ID Nº: 54 es la secuencia de ADNc determinada para 73264.
- 65 SEC ID Nº: 55 es la secuencia de ADNc determinada para 73266.
- SEC ID Nº: 56 es la secuencia de ADNc determinada para 73267.

- SEC ID N°: 57 es la secuencia de ADNc determinada para 73268.
- 5 SEC ID N°: 58 es la secuencia de ADNc determinada para 73269.
- SEC ID N°: 59 es la secuencia de ADNc determinada para 73270.
- SEC ID N°: 60 es la secuencia de ADNc determinada para 73272.
- 10 SEC ID N°: 61 es la secuencia de ADNc determinada para 73273.
- SEC ID N°: 62 es la secuencia de ADNc determinada para 73274.
- SEC ID N°: 63 es la secuencia de ADNc determinada para 73275.
- 15 SEC ID N°: 64 es la secuencia de ADNc determinada para 73037.
- SEC ID N°: 65 es la secuencia de ADNc determinada para 75038.
- 20 SEC ID N°: 66 es la secuencia de ADNc determinada para 75039.
- SEC ID N°: 67 es la secuencia de ADNc determinada para 75040.
- SEC ID N°: 68 es la secuencia de ADNc determinada para 75041.
- 25 SEC ID N°: 69 es la secuencia de ADNc determinada para 75042.
- SEC ID N°: 70 es la secuencia de ADNc determinada para 75044.
- 30 SEC ID N°: 71 es la secuencia de ADNc determinada para 75045.
- SEC ID N°: 72 es la secuencia de ADNc determinada para 75046.
- SEC ID N°: 73 es la secuencia de ADNc determinada para 75047.
- 35 SEC ID N°: 74 es la secuencia de ADNc determinada para 75048.
- SEC ID N°: 75 es la secuencia de ADNc de longitud completa determinada para 83537, también denominada flagelina X.
- 40 SEC ID N°: 76 es la secuencia de ADNc determinada para el extremo conservado amino-terminal de flagelina X.
- SEC ID N°: 77 es la secuencia de ADNc determinada para el extremo conservado amino-terminal más la región variable de flagelina X.
- 45 SEC ID N°: 78 es la secuencia de ADNc determinada para el extremo conservado carboxilo-terminal de flagelina X.
- SEC ID N°: 79 es la secuencia de aminoácidos de longitud completa de flagelina X.
- 50 SEC ID N°: 80 es la secuencia de aminoácidos del extremo conservado amino-terminal de flagelina X.
- SEC ID N°: 81 es la secuencia de aminoácidos del extremo conservado amino-terminal más la región variable de flagelina X.
- 55 SEC ID N°: 82 es la secuencia de aminoácidos del extremo conservado carboxilo-terminal de flagelina X.
- SEC ID N°: 83 es la secuencia codificante de longitud completa de flagelina B de *Helicobacter bilis*.
- 60 SEC ID N°: 84 es la secuencia de proteína de longitud completa de flagelina B de *Helicobacter bilis*, codificada por la secuencia de nucleótidos expuesta en SEC ID N°: 83.
- SEC ID N°: 85 es la secuencia codificante de longitud completa de flagelina de Cbir-1 (secuencia parcial expuesta en SEC ID N°: 1).
- 65 SEC ID N°: 86 es la secuencia de proteína de longitud completa de flagelina de Cbir-1, codificada por la secuencia de nucleótidos expuesta en SEC ID N°: 85.

**Descripción detallada de la invención**

5 La presente invención se refiere de manera general a composiciones y a su utilización en la terapia y el diagnóstico de EII. Tal como se describe adicionalmente a continuación, las composiciones ilustrativas de la presente invención incluyen, pero no se limitan a, polipéptidos, particularmente polipéptidos inmunogénicos, polinucleótidos que codifican para tales polipéptidos, anticuerpos y otros agentes de unión, células presentadoras de antígeno (CPA) y células del sistema inmunitario (por ejemplo, células T y B).

10 La práctica de la presente invención empleará, a menos que se indique específicamente lo contrario, procedimientos convencionales de virología, inmunología, microbiología, biología molecular y técnicas de ADN recombinante dentro de la experiencia de la técnica, muchos de los cuales se describen a continuación para fines de ilustración. Tales técnicas se explican completamente en la bibliografía. Véase, por ejemplo, Sambrook, *et al.* *Molecular Cloning: A Laboratory Manual* (2ª edición, 1989); Maniatis *et al.* *Molecular Cloning: A Laboratory Manual* (1982); *DNA Cloning: A Practical Approach*, vol. I y II (D. Glover, ed.); *Oligonucleotide Synthesis* (N. Gait, ed., 1984); *Nucleic Acid Hybridization* (B. Hames & S. Higgins, eds., 1985); *Transcription and Translation* (B. Hames & S. Higgins, eds., 1984); *Animal Cell Culture* (R. Freshney, ed., 1986); Perbal, *A Practical Guide to Molecular Cloning* (1984).

20 Tal como se utilizan en esta memoria descriptiva y en las reivindicaciones adjuntas, las formas singulares “un”, “una” y “el/la” incluyen referencias plurales a menos que el contenido indique claramente lo contrario.

Composiciones de polipéptido

25 Tal como se utiliza en la presente memoria, el término “polipéptido” se utiliza en su sentido convencional, es decir, como una secuencia de aminoácidos. Los polipéptidos no se limitan a una longitud específica del producto; por tanto, se incluyen péptidos, oligopéptidos y proteínas dentro de la definición de polipéptido, y tales términos pueden utilizarse de manera intercambiable en la presente memoria a menos que se indique específicamente lo contrario. Este término tampoco se refiere a, ni excluye, modificaciones tras la expresión del polipéptido, por ejemplo, glicosilaciones, acetilaciones, fosforilaciones y similares, así como otras modificaciones conocidas en la materia, tanto que se producen de manera natural como que no se producen de manera natural. Un polipéptido puede ser una proteína entera o una subsecuencia de la misma. Polipéptidos particulares de interés en el contexto de esta invención son subsecuencias de aminoácidos que comprenden epitopos, es decir, determinantes antigénicos sustancialmente responsables de las propiedades inmunogénicas de un polipéptido y que pueden provocar una respuesta inmunitaria.

35 Los polipéptidos ilustrativos particulares de la presente invención comprenden aquellos codificados por la secuencia de polinucleótido expuesta en SEC ID N°: 75. Algunos otros polipéptidos ilustrativos de la invención comprenden una secuencia de aminoácidos expuesta en SEC ID N°: 79.

40 Los polipéptidos de la presente invención se denominan algunas veces en la presente memoria proteínas bacterianas o polipéptidos bacterianos, como indicación de que su identificación se ha basado al menos parcialmente en su expresión en muestras de bacterias entéricas aisladas del colon de individuos con EII. Los péptidos descritos en la presente memoria pueden identificarse a partir de una lesión en el colon de un paciente con EII. Por consiguiente, un péptido de este tipo puede no estar presente en tejido normal adyacente. Alternativamente, un péptido de la presente invención puede identificarse a partir de una muestra de bacteria entérica aislada del colon de un individuo con EII, estando dicha bacteria entérica ausente en individuos no afectados por EII. Además, los polipéptidos de la presente invención pueden identificarse por su capacidad para activar células T a partir de individuos afectados por EII. Adicionalmente, pueden identificarse polipéptidos descritos en la presente memoria mediante su reactividad frente a sueros de pacientes con EII en comparación con su falta de reactividad frente a sueros de individuos no afectados.

55 Por tanto, un “polipéptido bacteriano” o “proteína bacteriana” se refiere de manera general a una secuencia de polipéptido de la presente invención, o una secuencia de polinucleótido que codifica para un polipéptido de este tipo, que está presente en muestras aisladas de una proporción sustancial de pacientes con EII, por ejemplo preferentemente superior a aproximadamente el 20%, más preferentemente superior a aproximadamente el 30% y lo más preferentemente superior a aproximadamente el 50% o más de pacientes sometidos a prueba según se determina utilizando un ensayo representativo proporcionado en la presente memoria. Una secuencia de polipéptido bacteriana de la invención, basándose en su expresión en muestras de bacterias entéricas aisladas del colon de individuos con EII, presenta utilidad particular tanto como marcador de diagnóstico así como diana terapéutica, tal como se describe adicionalmente a continuación. En una realización particular de la presente invención, un polipéptido bacteriano o proteína bacteriana comprende una proteína de flagelina.

60 En determinadas realizaciones preferidas, los polipéptidos de la invención son inmunogénicos, es decir, reaccionan de manera detectable dentro de un inmunoensayo (tal como un ensayo ELISA o de estimulación de células T) con antisueros y/o células T de un paciente con EII. Puede realizarse un examen para detectar actividad inmunogénica utilizando técnicas bien conocidas por el experto en la materia. Por ejemplo, tales exámenes pueden realizarse

utilizando procedimientos tales como los descritos en Harlow y Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. En un ejemplo ilustrativo, puede inmovilizarse un polipéptido sobre un soporte sólido y ponerse en contacto con sueros de pacientes para permitir la unión de anticuerpos dentro de los sueros con el polipéptido inmovilizado. Entonces pueden eliminarse los sueros no unidos y detectarse anticuerpos unidos utilizando, por ejemplo, proteína A marcada con <sup>125</sup>I.

Tal como reconocerá el experto en la materia, la presente invención también abarca partes inmunogénicas de los polipéptidos dados a conocer en la presente memoria. Una "parte inmunogénica", tal como se utiliza en la presente memoria, es un fragmento de un polipéptido inmunogénico de la invención que es en sí misma inmunológicamente reactiva (es decir, se une específicamente) con los receptores de antígeno de superficie de células B y/o células T que reconocen el polipéptido. Generalmente pueden identificarse partes inmunogénicas utilizando técnicas bien conocidas, tales como las resumidas en Paul, *Fundamental Immunology*, 3ª ed., 243-247 (Raven Press, 1993) y referencias mencionadas en el mismo. Tales técnicas incluyen examinar polipéptidos para detectar la capacidad para reaccionar con anticuerpos específicos, antisueros de antígeno y/o líneas o clones de células T. Tal como se utiliza en la presente memoria, los antisueros y anticuerpos son "específicos de antígeno" si se unen específicamente a un antígeno (es decir, reaccionan con la proteína en un ensayo ELISA u otro inmunoensayo, y no reaccionan de manera detectable con proteínas no relacionadas). Tales antisueros y anticuerpos pueden prepararse tal como se describe en la presente memoria y utilizando técnicas bien conocidas.

En una realización preferida, una parte inmunogénica de un polipéptido de la presente invención es una parte que reacciona con antisueros y/o células T a un nivel que no es sustancialmente inferior a la reactividad del polipéptido de longitud completa (por ejemplo, en un ensayo ELISA y/o de reactividad de células T). Preferentemente, el nivel de actividad inmunogénica de la parte inmunogénica es de al menos aproximadamente el 50%, preferentemente de al menos aproximadamente el 70% y lo más preferentemente superior a aproximadamente el 90% de la inmunogenicidad para el polipéptido de longitud completa. En algunos casos, se identificarán partes inmunogénicas preferidas que presentan un nivel de actividad inmunogénica superior al del polipéptido de longitud completa correspondiente, por ejemplo, que presentan más de aproximadamente el 100% o el 150% o más de actividad inmunogénica.

En determinadas otras realizaciones, las partes inmunogénicas ilustrativas pueden incluir péptidos en los que se ha deletado una secuencia líder N-terminal y/o dominio transmembrana. Otras partes inmunogénicas ilustrativas contendrán una pequeña delección N y/o C-terminal (por ejemplo, 1-30 aminoácidos, preferentemente 5-15 aminoácidos), con respecto a la proteína madura.

En otra realización, una composición de polipéptido de la invención también puede comprender uno o más polipéptidos que son inmunológicamente reactivos con células T y/o anticuerpos generados contra un polipéptido de la invención, particularmente un polipéptido que presenta una secuencia de aminoácidos dada a conocer en la presente memoria, o con un fragmento o variante inmunogénico del mismo.

En otra realización de la invención, se proporcionan polipéptidos que comprenden uno o más polipéptidos que pueden provocar células T y/o anticuerpos que son inmunológicamente reactivos con uno o más polipéptidos descritos en la presente memoria, o uno o más polipéptidos codificados por secuencias de ácido nucleico contiguas contenidas en las secuencias de polinucleótido dadas a conocer en la presente memoria, o fragmentos o variantes inmunogénicos de los mismos, o con una o más secuencias de ácido nucleico que se hibridan con una o más de esas secuencias en condiciones de rigurosidad de moderada a alta.

La presente invención, en otro aspecto, proporciona fragmentos de polipéptido que comprenden al menos aproximadamente 25, 50 o 100 aminoácidos contiguos, o más, incluyendo todas las longitudes intermedias de SEC ID N°: 79

En otro aspecto, la presente invención proporciona variantes de las composiciones de polipéptido descritas en la presente memoria. Las variantes de polipéptido generalmente abarcadas por la presente invención mostrarán normalmente al menos aproximadamente el 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% o el 99% o más de identidad (determinada tal como se describe a continuación), a lo largo de su longitud, con la secuencia expuesta en SEC ID N°: 79.

Los fragmentos de polipéptido proporcionados por la presente invención son inmunológicamente reactivos con un anticuerpo que reacciona con el polipéptido de longitud completa de SEC ID N°: 79.

En otra realización preferida, los fragmentos y variantes de polipéptido proporcionados por la presente invención muestran un nivel de actividad inmunogénica de al menos aproximadamente el 50%, preferentemente al menos aproximadamente el 70% y lo más preferentemente al menos el aproximadamente 90% o más de la mostrada por una secuencia de polipéptido de longitud completa específicamente expuesta en la presente memoria.

Una "variante" de polipéptido, tal como se utiliza el término en la presente memoria, es un polipéptido que se diferencia normalmente de un polipéptido específicamente dado a conocer en la presente memoria en una o más

sustituciones, deleciones, adiciones y/o inserciones. Tales variantes pueden producirse de manera natural o pueden generarse de manera sintética, por ejemplo, modificando una o más de las secuencias de polipéptido anteriores de la invención y evaluando su actividad inmunogénica tal como se describe en la presente memoria y/o utilizando cualquiera de varias técnicas bien conocidas en la materia.

Por ejemplo, determinadas variantes ilustrativas de los polipéptidos de la invención incluyen aquellas en las que se han eliminado una o más partes, tales como una secuencia líder N-terminal o dominio transmembrana. Otras variantes ilustrativas incluyen variantes en las que se ha eliminado una pequeña parte (por ejemplo, 1-30 aminoácidos, preferentemente 5-15 aminoácidos) del extremo N y/o C-terminal de la proteína madura.

En muchos casos, una variante contendrá sustituciones conservativas. Una "sustitución conservativa" es una en la que se sustituye un aminoácido por otro aminoácido que presenta propiedades similares, de tal manera que un experto en la materia de química de péptidos esperará que la estructura secundaria y la naturaleza hidropática del polipéptido permanezcan sustancialmente inalteradas. Tal como se describió anteriormente, pueden realizarse modificaciones en la estructura de los polinucleótidos y polipéptidos de la presente invención y todavía obtener una molécula funcional que codifica para un polipéptido variante o derivado con características deseables, por ejemplo, con características inmunogénicas. Cuando se desea alterar la secuencia de aminoácidos de un polipéptido para crear una variante o parte inmunogénica equivalente, o incluso mejorada, de un polipéptido de la invención, un experto en la materia cambiará normalmente uno o más de los codones de la secuencia de ADN codificante según la tabla 1.

Por ejemplo, determinados aminoácidos pueden sustituirse por otros aminoácidos en una estructura proteica sin pérdida apreciable de capacidad de unión interactiva con estructuras tales como, por ejemplo, regiones de unión a antígeno de anticuerpos o sitios de unión en moléculas de sustrato. Dado que es la capacidad interactiva y la naturaleza de una proteína lo que define la actividad funcional biológica de la proteína, pueden realizarse determinadas sustituciones de secuencia de aminoácidos en una secuencia de proteína, y, evidentemente, su secuencia codificante de ADN subyacente, y no obstante obtener una proteína con propiedades similares. Por tanto, se contempla que pueden realizarse diversos cambios en las secuencias peptídicas de las composiciones dadas a conocer, o secuencias de ADN correspondientes que codifican para dichos péptidos, sin pérdida apreciable de su actividad o utilidad biológica.

TABLA 1

Aminoácidos		Codones	
Alanina	Ala	A	GCA GCC GCG GCU
Cisteína	Cys	C	UGC UGU
Ácido aspártico	Asp	D	GAC GAU
Ácido glutámico	Glu	E	GAA GAG
Fenilalanina	Phe	F	UUC UUU
Glicina	Gly	G	GGA GGC GGG GGU
Histidina	His	H	CAC CAU
Isoleucina	Ile	I	AUA AUC AUU
Lisina	Lys	K	AAA AAG
Leucina	Leu	L	UUA UUG CUA CUC CUG CUU
Metionina	Met	M	AUG
Asparagina	Asn	N	AAC AAU
Prolina	Pro	P	CCA CCC CCG CCU
Glutamina	Gln	Q	CAA CAG
Arginina	Arg	R	AGA AGG CGA CGC CGG CGU
Serina	Ser	S	AGC AGU UCA UCC UCG UCU
Treonina	Thr	T	ACA ACC ACG ACU
Valina	Val	V	GUA GUC GUG GUU
Triptófano	Trp	W	UGG
Tirosina	Tyr	Y	UAC UAU

En la realización de tales cambios, puede considerarse el índice hidropático de los aminoácidos. La importancia del índice hidropático del aminoácido para conferir función biológica interactiva a una proteína se entiende generalmente en la materia (Kyte y Doolittle, 1982). Se acepta que el carácter hidropático relativo del aminoácido contribuye a la estructura secundaria de la proteína resultante, que a su vez define la interacción de la proteína con otras moléculas, por ejemplo, enzimas, sustratos, receptores, ADN, anticuerpos, antígenos y similares. A cada aminoácido se le ha asignado un índice hidropático basándose en su hidrofobia y características de carga (Kyte y Doolittle, 1982). Estos valores son: isoleucina (+4,5); valina (+4,2); leucina (+3,8); fenilalanina (+2,8); cisteína/cistina (+2,5); metionina (+1,9); alanina (+1,8); glicina (-0,4); treonina (-0,7); serina (-0,8); triptófano (-0,9); tirosina (-1,3); prolina (-1,6); histidina (-3,2); glutamato (-3,5); glutamina (-3,5); aspartato (-3,5); asparagina (-3,5); lisina (-3,9) y arginina (-4,5).

En la materia se sabe que determinados aminoácidos pueden sustituirse por otros aminoácidos que presentan una

puntuación o índice hidropático similar y todavía dar como resultado una proteína con actividad biológica similar, es decir, todavía obtener una proteína biológica funcionalmente equivalente. Al realizar tales cambios, se prefiere la sustitución de aminoácidos cuyos índices hidropáticos están dentro de  $\pm 2$ , se prefieren particularmente aquellos dentro de  $\pm 1$  y se prefieren incluso más particularmente aquellos dentro de  $\pm 0,5$ . También se entiende en la materia

5 que la sustitución de aminoácidos similares puede realizarse eficazmente basándose en su hidrofilia. La patente US 4.554.101 menciona que la mayor hidrofilia promedio local de una proteína, regida por la hidrofilia de sus aminoácidos adyacentes, se correlaciona con una propiedad biológica de la proteína.

Tal como detalla la patente US 4.554.101, se han asignado los siguientes valores de hidrofilia a residuos de

10 aminoácido: arginina (+3,0); lisina (+3,0); aspartato (+3,0  $\pm$  1); glutamato (+3,0  $\pm$  1); serina (+0,3); asparagina (+0,2); glutamina (+0,2); glicina (0); treonina (-0,4); prolina (-0,5  $\pm$  1); alanina (-0,5); histidina (-0,5); cisteína (-1,0); metionina (-1,3); valina (-1,5); leucina (-1,8); isoleucina (-1,8); tirosina (-2,3); fenilalanina (-2,5); triptófano (-3,4). Se entiende que puede sustituirse un aminoácido por otro que presenta un valor de hidrofilia similar y todavía obtener una proteína biológicamente equivalente, y en particular, inmunológicamente equivalente. En tales cambios, se prefiere

15 la sustitución de aminoácidos cuyos valores de hidrofilia están dentro de  $\pm 2$ , se prefieren particularmente aquellos dentro de  $\pm 1$  y se prefieren incluso más particularmente aquellos dentro de  $\pm 0,5$ .

Tal como se expuso anteriormente, las sustituciones de aminoácidos se basan por tanto generalmente en la similitud relativa de los sustituyentes de cadena lateral de los aminoácidos, por ejemplo, su hidrofobia, hidrofilia, carga, tamaño y similares. Las sustituciones a modo de ejemplo que presentan varias de las características anteriores en cuenta las conocen bien los expertos en la materia e incluyen: arginina y lisina; glutamato y aspartato; serina y treonina; glutamina y asparagina; y valina, leucina e isoleucina.

20

Además, puede modificarse adicionalmente cualquier polinucleótido para aumentar la estabilidad *in vivo*. Las posibles modificaciones incluyen, pero no se limitan a, la adición de secuencias flanqueantes en los extremos 5' y/o 3'; la utilización de fosforotioato o 2'-O-metilo en vez de enlaces fosfodiesterasa en la estructura principal; y/o la inclusión de bases no tradicionales tales como inosina, queosina y wybutosina, así como formas modificadas con acetilo, metilo, tio y otras de adenina, citidina, guanina, timina y uridina.

25

Pueden realizarse adicionalmente sustituciones de aminoácidos basándose en la similitud de polaridad, carga, solubilidad, hidrofobia, hidrofilia y/o la naturaleza anfipática de los residuos. Por ejemplo, los aminoácidos cargados negativamente incluyen ácido aspártico y ácido glutámico; los aminoácidos cargados positivamente incluyen lisina y arginina; y los aminoácidos con grupos de cabeza polares no cargados que presentan valores de hidrofilia similares incluyen leucina, isoleucina y valina; glicina y alanina; asparagina y glutamina; y serina, treonina, fenilalanina y tirosina. Otros grupos de aminoácidos que pueden representar cambios conservativos incluyen: (1) ala, pro, gly, glu, asp, gln, asn, ser, thr; (2) cys, ser, tyr, thr; (3) val, ile, leu, met, ala, phe; (4) lys, arg, his y (5) phe, tyr, trp, his. Una variante también, o alternativamente, puede contener cambios no conservativos. En una realización preferida, polipéptidos variantes se diferencian de una secuencia nativa mediante sustitución, delección o adición de cinco aminoácidos o menos. También (o alternativamente) pueden modificarse variantes mediante, por ejemplo, la delección o adición de aminoácidos que presentan una influencia mínima sobre la inmunogenicidad, estructura secundaria y naturaleza hidropática del polipéptido.

30

35

40

Tal como se indicó anteriormente, los polipéptidos pueden comprender una secuencia señal (o líder) en el extremo N-terminal de la proteína, que de manera cotransduccional o postransduccional dirige la transferencia de la proteína. El polipéptido también puede conjugarse con un ligador u otra secuencia por facilidad de síntesis, purificación o identificación del polipéptido (por ejemplo, poli-His), o para potenciar la unión del polipéptido a un soporte sólido. Por ejemplo, un polipéptido puede conjugarse con una región Fc de inmunoglobulina.

45

Cuando se comparan secuencias de polipéptido, se dice que dos secuencias son "idénticas" si la secuencia de aminoácidos en las dos secuencias es la misma cuando se alinean para obtener máxima correspondencia, tal como se describe a continuación. Normalmente se realizan comparaciones entre dos secuencias comparando las secuencias a lo largo de una ventana de comparación para identificar y comparar regiones locales de similitud de secuencia. Una "ventana de comparación" tal como se utiliza en la presente memoria, se refiere a un segmento de al menos aproximadamente 20 posiciones contiguas, habitualmente de 30 a aproximadamente 75, de 40 a

50

55 aproximadamente 50, en la que puede compararse una secuencia con una secuencia de referencia con el mismo número de posiciones contiguas tras alinear óptimamente las dos secuencias.

Puede realizarse la alineación óptima de secuencias para su comparación utilizando el programa Megalign en la serie Lasergene de software de bioinformática (DNASTAR, Inc., Madison, WI), utilizando los parámetros por defecto. Este programa realiza varios esquemas de alineación descritos en las siguientes referencias: Dayhoff, M.O. (1978) A model of evolutionary change in proteins - Matrices for detecting distant relationships. En Dayhoff, M.O. (ed.) Atlas of Protein Sequence and Structure, National Biomedical Research Foundation, Washington DC Vol. 5, Sup. 3, págs. 345-358; Hein J. (1990) Unified Approach to Alignment and Phylogenies págs. 626-645 Methods in Enzymology vol. 183, Academic Press, Inc., San Diego, CA; Higgins, D.G. y Sharp, P.M. (1989) CABIOS 5:151-153; Myers, E.W. y Muller W. (1988) CABIOS 4:11-17; Robinson, E.D. (1971) Comb. Theor 11:105; Saitou, N. Nei, M. (1987) Mol. Biol. Evol. 4:406-425; Sneath, P.H.A. y Sokal, R.R. (1973) Numerical Taxonomy - the Principles and Practice of Numerical

60

65

Taxonomy, Freeman Press, San Francisco, CA; Wilbur, W.J. y Lipman, D.J. (1983) Proc. Natl. Acad., Sci. USA 80:726-730.

Alternativamente, puede realizarse la alineación óptima de secuencias para su comparación mediante el algoritmo de identidad local de Smith y Waterman (1981) Add. APL. Math 2:482, mediante el algoritmo de alineación de identidad de Needleman y Wunsch (1970) J. Mol. Biol. 48:443, mediante los procedimientos de búsqueda de similitud de Pearson y Lipman (1988) Proc. Natl. Acad. Sci. USA 85: 2444, mediante implementaciones informáticas de estos algoritmos (GAP, BESTFIT, BLAST, FASTA, y TFASTA en el paquete de software de Wisconsin Genetics, Genetics Computer Group (GCG), 575 Science Dr., Madison, WI) o mediante inspección.

Un ejemplo preferido de algoritmos que son adecuados para determinar el porcentaje de identidad de secuencia y similitud de secuencia son los algoritmos BLAST y BLAST 2.0, que se describen en Altschul *et al.* (1977) Nucl. Acids Res. 25: 3389-3402 y Altschul *et al.* (1990) J. Mol. Biol. 215:403-410, respectivamente. BLAST y BLAST 2.0 pueden utilizarse, por ejemplo, con los parámetros descritos en la presente memoria, para determinar el porcentaje de identidad de secuencia para los polinucleótidos y polipéptidos de la invención. El software para realizar análisis BLAST está disponible al público a través del National Center for Biotechnology Information. Para secuencias de aminoácidos, puede utilizarse una matriz de puntuación para calcular la puntuación acumulativa. La extensión de los resultados positivos de palabra se detiene cuando: la puntuación de alineación acumulativa disminuye en una cantidad X desde su valor máximo alcanzado; la puntuación acumulativa llega hasta cero o menos, debido a la acumulación de una o más alineaciones de residuos con puntuación negativa; o se alcanza el final de cualquier secuencia. Los parámetros del algoritmo BLAST, W, T y X, determinan la sensibilidad y velocidad de la alineación.

En un enfoque preferido, el "porcentaje de identidad de secuencia" se determina comparando dos secuencias alineadas óptimamente a lo largo de una ventana de comparación de al menos 20 posiciones, en la que la parte de la secuencia de polipéptido en la ventana de comparación puede comprender adiciones o deleciones (es decir, huecos) del 20 por ciento o menos, habitualmente del 5 al 15 por ciento, o del 10 al 12 por ciento, en comparación con las secuencias de referencia (que no comprenden adiciones o deleciones) para una alineación óptima de las dos secuencias. El porcentaje se calcula determinando el número de posiciones en las que se produce el residuo de aminoácido idéntico en ambas secuencias para proporcionar el número de posiciones coincidentes, dividiendo el número de posiciones coincidentes entre el número total de posiciones en la secuencia de referencia (es decir, el tamaño de ventana) y multiplicando los resultados por 100 para proporcionar el porcentaje de identidad de secuencia.

Dentro de otras realizaciones ilustrativas, un polipéptido puede ser un polipéptido de fusión que comprende múltiples polipéptidos tal como se describe en la presente memoria, o que comprende al menos un polipéptido tal como se describe en la presente memoria y una secuencia no relacionada, tal como una proteína bacteriana conocida. Una pareja de fusión puede, por ejemplo, ayudar a proporcionar epítopos de células cooperadoras T (una pareja de fusión inmunológica), preferentemente epítopos de células cooperadoras T reconocidos por seres humanos, o pueden ayudar a expresar la proteína (un potenciador de la expresión) a rendimientos superiores a los de la proteína recombinante nativa. Determinadas parejas de fusión preferidas son tanto parejas de fusión inmunológicas como que potencian la expresión. Otras parejas de fusión pueden seleccionarse de modo que se aumenta la solubilidad del polipéptido o se permite dirigir el polipéptido a compartimentos intracelulares deseados. Parejas de fusión todavía adicionales incluyen etiquetas de afinidad, que facilitan la purificación del polipéptido.

Los polipéptidos de fusión pueden prepararse generalmente utilizando técnicas convencionales, incluyendo conjugación química. Preferentemente, un polipéptido de fusión se expresa como polipéptido recombinante, permitiendo la producción de niveles aumentados, con respecto a un polipéptido no fusionado, en un sistema de expresión. En resumen, pueden ensamblarse por separado secuencias de ADN que codifican para los componentes del polipéptido, y ligarse en un vector de expresión apropiado. Se liga el extremo 3' de la secuencia de ADN que codifica para un componente de polipéptido, con o sin un ligador peptídico, al extremo 5' de una secuencia de ADN que codifica para el segundo componente de polipéptido de modo que los marcos de lectura de las secuencias están en fase. Esto permite la traducción en un único polipéptido de fusión que conserva la actividad biológica de ambos polipéptidos componentes.

Puede emplearse una secuencia de ligador peptídico para separar los componentes de polipéptido primero y segundo una distancia suficiente para garantizar que cada polipéptido se pliega en sus estructuras secundaria y terciaria. Una secuencia de ligador peptídico de este tipo se incorpora en el polipéptido de fusión utilizando técnicas convencionales bien conocidas en la materia. Pueden elegirse secuencias de ligador peptídico adecuadas basándose en los siguientes factores: (1) su capacidad para adoptar una conformación extendida flexible; (2) su incapacidad para adoptar una estructura secundaria que pudiera interactuar con epítopos funcionales en los polipéptidos primero y segundo; y (3) la falta de residuos hidrófobos o cargados que puedan reaccionar con los epítopos funcionales del polipéptido. Las secuencias de ligador peptídico preferidas contienen residuos de Gly, Asn y Ser. También pueden utilizarse otros aminoácidos casi neutros, tales como Thr y Ala, en la secuencia de ligador. Las secuencias de aminoácidos que pueden emplearse de manera útil como ligadores incluyen las dadas a conocer en Maratea *et al.*, Gene 40:39-46, 1985; Murphy *et al.*, Proc. Natl. Acad. Sci. USA 83:8258-8262, 1986; la patente US n.º 4.935.233 y la patente US n.º 4.751.180. La secuencia de ligador puede presentar generalmente desde 1

hasta aproximadamente 50 aminoácidos de longitud. Las secuencias de ligador no se requieren cuando los polipéptidos primero y segundo presentan regiones de aminoácidos N-terminales no esenciales que pueden utilizarse para separar los dominios funcionales y evitar la interferencia estérica.

5 Las secuencias de ADN ligadas están funcionalmente unidas a elementos reguladores de la transcripción o de la traducción adecuados. Los elementos reguladores responsables de la expresión de ADN sólo están situados en sentido 5' con respecto a la secuencia de ADN que codifica para los primeros polipéptidos. De manera similar, los codones de terminación requeridos para terminar la traducción las señales de terminación de la transcripción sólo están presentes en sentido 3' con respecto a la secuencia de ADN que codifica para el segundo polipéptido.

10 El polipéptido de fusión puede comprender un polipéptido tal como se describe en la presente memoria junto con una proteína inmunogénica no relacionada, tal como una proteína inmunogénica que puede provocar una respuesta de recuerdo. Los ejemplos de tales proteínas incluyen proteínas de tétanos, tuberculosis y hepatitis (véase, por ejemplo, Stoute *et al.* New Engl. J. Med., 336:86-91, 1997).

15 En una realización preferida, la pareja de fusión inmunológica se deriva de una *Mycobacterium sp.*, tal como un fragmento de Ra12 derivado de *Mycobacterium tuberculosis*. En la solicitud de patente US 60/158.585 se describen composiciones de Ra12 y procedimientos para su utilización en la potenciación de la expresión y/o la inmunogenicidad de secuencias de polinucleótido/polipéptido heterólogas. En resumen, Ra12 se refiere a una región de polinucleótido que es una subsecuencia de un ácido nucleico de MTB32A de *Mycobacterium tuberculosis*.  
 20 MTB32A es una serina proteasa con un peso molecular de 32 KD codificada por un gen en cepas virulentas y avirulentas de *M. tuberculosis*. La secuencia de nucleótidos y la secuencia de aminoácidos de MTB32A se han descrito (por ejemplo, solicitud de patente US 60/158.585; véase también, Skeiky *et al.*, Infection and Immun. (1999) 67: 3998-4007). Fragmentos C-terminales de la secuencia que codifica para MTB32A se expresan a altos niveles y permanecen como polipéptidos solubles a lo largo de todo el procedimiento de purificación. Además, Ra12 puede potenciar la inmunogenicidad de polipéptidos inmunogénicos heterólogos con los que se fusiona. Un polipéptido de fusión de Ra12 preferido comprende un fragmento C-terminal de 14 KD correspondiente a los residuos de aminoácido de 192 a 323 de MTB32A. Otros polinucleótidos de Ra12 preferidos comprenden generalmente al menos aproximadamente 15 nucleótidos consecutivos, al menos aproximadamente 30 nucleótidos, al menos aproximadamente 60 nucleótidos, al menos aproximadamente 100 nucleótidos, al menos aproximadamente 200 nucleótidos o al menos aproximadamente 300 nucleótidos que codifican para una parte de un polipéptido Ra12. Los polinucleótidos de Ra12 pueden comprender una secuencia nativa (es decir, una secuencia endógena que codifica para un polipéptido Ra12 o una parte del mismo) o pueden comprender una variante de una secuencia de este tipo. Las variantes de polinucleótido de Ra12 pueden contener una o más sustituciones, adiciones, deleciones y/o inserciones de tal manera que la actividad biológica del polipéptido de fusión codificado no se ve sustancialmente disminuida, con respecto a un polipéptido de fusión que comprende un polipéptido Ra12 nativo. Las variantes muestran preferentemente al menos aproximadamente el 70% de identidad, más preferentemente al menos aproximadamente el 80% de identidad y lo más preferentemente al menos aproximadamente el 90% de identidad con una secuencia de polinucleótido que codifica para un polipéptido Ra12 nativo o una parte del mismo.

40 Dentro de otras realizaciones preferidas, una pareja de fusión inmunológica se deriva de proteína D, una proteína de superficie de la bacteria gram-negativa *Haemophilus influenza B* (documento WO 91/18926). Preferentemente, un derivado de proteína D comprende aproximadamente el primer tercio de la proteína (por ejemplo, los primeros 100-110 aminoácidos N-terminales), y un derivado de proteína D puede estar lipidado. Dentro de determinadas realizaciones preferidas, se incluyen los primeros 109 residuos de una pareja de fusión de lipoproteína D en el extremo N-terminal para proporcionar al polipéptido epítopos de células T exógenos adicionales y para aumentar el nivel de expresión en *E. coli* (funcionando por tanto como potenciador de expresión). La cola de lípido garantiza la presentación óptima del antígeno a células presentadoras de antígeno. Otras parejas de fusión incluyen la proteína no estructural de virus influenza, NS1 (hemaglutinina). Normalmente, se utilizan los 81 aminoácidos N-terminales, aunque pueden utilizarse diferentes fragmentos que incluyen epítopos de células cooperadoras T.

55 En otra realización, la pareja de fusión inmunológica es la proteína conocida como LYTA o una parte de la misma (preferentemente una parte C-terminal). LYTA se deriva de *Streptococcus pneumoniae*, que sintetiza una N-acetil-L-alanina amidasa conocida como amidasa LYTA (codificada por el gen *LytA*; Gene 43:265-292, 1986). LYTA es una autolisina que degrada específicamente determinados enlaces en la estructura principal de peptidoglicanos. El dominio C-terminal de la proteína LYTA es responsable de la afinidad a la colina o a algunos análogos de colina tales como DEAE. Esta propiedad se ha aprovechado para el desarrollo de plásmidos que expresan C-LYTA de *E. coli* útiles para la expresión de proteínas de fusión. Se ha descrito la purificación de proteínas híbridas que contienen el fragmento C-LYTA en el extremo amino-terminal (véase Biotechnology 10:795-798, 1992). Dentro de una realización preferida, puede incorporarse una parte de repetición de LYTA en un polipéptido de fusión. Se encuentra una parte de repetición en la región C-terminal comenzando en el residuo 178. Una parte de repetición particularmente preferida incorpora los residuos 188-305.

65 Aún otra realización ilustrativa implica polipéptidos de fusión, y los polinucleótidos que codifican para los mismos, en los que la pareja de fusión comprende una señal de direccionamiento que puede dirigir un polipéptido al compartimento endosómico/lisosómico, tal como se describe en la patente US n.º 5.633.234. Un polipéptido

inmunogénico de la invención, cuando se fusiona con esta señal de direccionamiento, se asociará más eficazmente con moléculas de la clase II del MHC y proporcionará así estimulación *in vivo* potenciada de células T CD4+ específicas para el polipéptido.

5 Se preparan polipéptidos de la invención utilizando cualquiera de una variedad de técnicas sintéticas y/o recombinantes bien conocidas, las últimas de las cuales se describen adicionalmente a continuación. Pueden generarse polipéptidos, partes y otras variantes de generalmente menos de aproximadamente 150 aminoácidos mediante medios sintéticos, utilizando técnicas bien conocidas por los expertos habituales en la materia. En un ejemplo ilustrativo, tales polipéptidos se sintetizan utilizando cualquiera de las técnicas en fase sólida disponibles comercialmente, tales como el procedimiento de síntesis en fase sólida de Merrifield, en el que se añaden aminoácidos secuencialmente a una cadena de aminoácido en crecimiento. Véase Merrifield, J. Am. Chem. Soc. 85:2149-2146, 1963. El equipo para la síntesis automática de polipéptidos está disponible comercialmente de proveedores tales como Perkin Elmer/Applied BioSystems Division (Foster City, CA) y puede hacerse funcionar según las instrucciones del fabricante.

15 En general, las composiciones de polipéptido (incluyendo polipéptidos de fusión) de la invención están aisladas. Un polipéptido "aislado" es uno que se retira de su entorno original. Por ejemplo, una proteína o polipéptido que se produce de manera natural está aislado si está separado de algunos o todos los materiales coexistentes en el sistema natural. Preferentemente, tales polipéptidos también están purificados, por ejemplo, son puros en al menos aproximadamente el 90%, más preferentemente puros en al menos aproximadamente el 95% y lo más preferentemente puros en al menos aproximadamente el 99%.

#### Composiciones de polinucleótido

25 La presente invención, en otros aspectos, proporciona composiciones de polinucleótido. Los términos "ADN" y "polinucleótido" se utilizan esencialmente de manera intercambiable en la presente memoria para referirse a una molécula de ADN que se ha aislado libre de ADN genómico total de una especie particular. "Aislado", tal como se utiliza en la presente memoria, significa que un polinucleótido está sustancialmente lejos de otras secuencias codificantes, y que la molécula de ADN no contiene partes grandes de ADN codificante no relacionado, tal como fragmentos cromosómicos grandes u otros genes funcionales o regiones codificantes de polipéptido. Evidentemente, se refiere a la molécula de ADN según se aísla originalmente, y no excluye genes o regiones codificantes posteriormente añadidas al segmento manualmente.

35 Tal como entenderán los expertos en la materia, las composiciones de polinucleótido de esta invención pueden incluir secuencias genómicas, secuencias extragenómicas y codificadas por plásmido y segmentos génicos diseñados mediante ingeniería más pequeños que expresan, o pueden adaptarse para expresar, proteínas, polipéptidos, péptidos y similares. Tales segmentos pueden estar aislados de manera natural o modificarse sintéticamente de manera manual.

40 Tal como reconocerá el experto en la materia, los polinucleótidos de la invención pueden ser monocatenarios (codificantes o antisentido) o bicatenarios, y pueden ser moléculas de ADN (genómico, ADNc o sintético) o de ARN. Las moléculas de ARN pueden incluir moléculas de ARNnH, que contiene intrones y corresponde a una molécula de ADN de una manera uno a uno, y moléculas de ARNm, que no contiene intrones. Pueden estar presentes, pero no son necesarias, secuencias codificantes o no codificantes adicionales dentro de un polinucleótido de la presente invención, y un polinucleótido puede estar unido, pero no necesariamente, a otras moléculas y/o materiales de soporte.

50 Los polinucleótidos pueden comprender una secuencia nativa (es decir, una secuencia endógena que codifica para un polipéptido/proteína de la invención o una parte del mismo) o pueden comprender una secuencia que codifica para una variante o derivado, preferentemente una variante o derivado inmunogénico, de una secuencia de este tipo.

55 Por tanto, según otro aspecto de la presente invención, se proporcionan composiciones de polinucleótido que comprenden la secuencia de polinucleótido expuesta en SEC ID N°: 75, el complemento de la secuencia de polinucleótido expuesta en SEC ID N°: 75, y variantes degeneradas de la secuencia de polinucleótido expuesta en SEC ID N°: 75. En determinadas realizaciones preferidas, las secuencias de polinucleótido expuestas en la presente memoria codifican para polipéptidos inmunogénicos, tal como se describió anteriormente. En otras determinadas realizaciones preferidas, las secuencias de polinucleótido expuestas en la presente memoria codifican para proteínas bacterianas asociadas con EII aisladas tal como se describe en la presente memoria a partir de individuos afectados por EII. En otras determinadas realizaciones preferidas, las secuencias de polinucleótido expuestas en la presente memoria codifican para proteínas de flagelina.

65 En otras realizaciones relacionadas, la presente invención proporciona variantes de polinucleótido que presentan identidad sustancial con la secuencia dada a conocer en la presente memoria en SEC ID N°: 75, por ejemplo aquellas que comprende al menos el 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% o el 99% o más de identidad de secuencia en comparación con la secuencia de SEC ID N°: 75 utilizando los procedimientos descritos en la presente

memoria (por ejemplo, análisis de BLAST utilizando parámetros convencionales, tal como se describe a continuación). Un experto en la materia reconocerá que estos valores pueden ajustarse de manera apropiada para determinar la identidad correspondiente de proteínas codificadas por dos secuencias de nucleótidos teniendo en cuenta la degeneración de codones, similitud de aminoácidos, colocación del marco de lectura y similares.

Normalmente, las variantes de polinucleótido contendrán una o más sustituciones, adiciones, deleciones y/o inserciones, preferentemente de tal manera que la inmunogenicidad del polipéptido codificado por el polinucleótido variante no se ve sustancialmente disminuida con respecto a un polipéptido codificado por una secuencia de polinucleótido específicamente expuesta en la presente memoria). También debe entenderse que el término "variantes" abarca genes homólogos de origen xenogénico.

También se dan a conocer fragmentos de polinucleótido que comprenden o consisten en diversas longitudes de tramos contiguos de secuencia idéntica o complementaria a una o más de las secuencias dadas a conocer en la presente memoria. Por ejemplo, se dan a conocer polinucleótidos que comprenden o consisten en al menos aproximadamente 10, 15, 20, 30, 40, 50, 75, 100, 150, 200, 300, 400, 500 o 1000 o más nucleótidos contiguos de una o más de las secuencias dadas a conocer en la presente memoria así como todas las longitudes intermedias entre las mismas. Se entenderá fácilmente que "longitudes intermedias", en este contexto, significa cualquier longitud entre los valores mencionados, tales como 16, 17, 18, 19, etc.; 21, 22, 23, etc.; 30, 31, 32, etc.; 50, 51, 52, 53, etc.; 100, 101, 102, 103, etc.; 150, 151, 152, 153, etc.; incluyendo todos los enteros hasta 200-500; 500-1.000 y similares. Una secuencia de polinucleótido tal como se describe en la presente memoria puede extenderse en uno o ambos extremos mediante nucleótidos adicionales no encontrados en la secuencia nativa. Esta secuencia adicional puede consistir en 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 o 20 nucleótidos en cualquier extremo de la secuencia dada a conocer o en ambos extremos de la secuencia dada a conocer.

También se dan a conocer composiciones de polinucleótido que pueden hibridarse en condiciones de rigurosidad de moderada a alta con una secuencia de polinucleótido proporcionada en la presente memoria, o un fragmento de la misma, o una secuencia complementaria de la misma. Las técnicas de hibridación se conocen bien en la materia de biología molecular. Para fines de ilustración, las condiciones de rigurosidad moderada adecuadas para someter a prueba la hibridación de un polinucleótido de esta invención con otros polinucleótidos incluyen lavado previo en una disolución de 5 X SSC, el 0,5% de SDS, EDTA 1,0 mM (pH 8,0); hibridación a 50gC-609.C, 5 X SSC, durante la noche; seguido por lavado dos veces a 65-9C durante 20 minutos con cada uno de 2X, 0,5X y 0,2X SSC que contiene el 0,1% de SDS. Un experto en la materia entenderá que la rigurosidad de hibridación puede manipularse fácilmente, tal como alterando el contenido en sal de la disolución de hibridación y/o la temperatura a la cual se realiza la hibridación. Por ejemplo, las condiciones de hibridación de alta rigurosidad adecuadas incluyen las descritas anteriormente, con la excepción de que se aumenta la temperatura de hibridación, por ejemplo, hasta 60-65°C o 65-70°C.

Los polinucleótidos descritos anteriormente, por ejemplo, variantes, fragmentos y secuencias de hibridación de polinucleótido, pueden codificar para polipéptidos que reaccionan de manera cruzada inmunológicamente con una secuencia de polipéptido específicamente expuesta en la presente memoria. Tales polinucleótidos pueden codificar para polipéptidos que presentan un nivel de actividad inmunogénica de al menos aproximadamente el 50%, preferentemente de al menos aproximadamente el 70% y más preferentemente de al menos aproximadamente el 90% del de una secuencia de polipéptido específicamente expuesta en la presente memoria.

Los polinucleótidos de la presente invención, o fragmentos de los mismos, independientemente de la longitud de la propia secuencia codificante, pueden combinarse con otras secuencias de ADN, tales como promotores, señales de poliadenilación, sitios de enzima de restricción adicionales, sitios de clonación múltiple, otros segmentos codificantes, y similares, de tal manera que su longitud global puede variar considerablemente. Por tanto se contempla que puede emplearse un fragmento de ácido nucleico de casi cualquier longitud, limitándose la longitud total preferentemente por la facilidad de preparación y utilización en el protocolo de ADN recombinante pretendido. Por ejemplo, se contempla que segmentos de polinucleótido ilustrativos con longitudes totales de aproximadamente 10.000, aproximadamente 5000, aproximadamente 3000, aproximadamente 2.000, aproximadamente 1.000, aproximadamente 500, aproximadamente 200, aproximadamente 100, aproximadamente 50 pares de bases de longitud, y similares (incluyendo todas las longitudes intermedias) son útiles en muchas implementaciones de esta invención.

Cuando se comparan secuencias de polinucleótido, se dice que dos secuencias son "idénticas" si la secuencia de nucleótidos en las dos secuencias es la misma cuando se alinean para su máxima correspondencia, tal como se describe a continuación. Las comparaciones entre dos secuencias se realizan normalmente comparando las secuencias a lo largo de una ventana de comparación para identificar y comparar regiones locales de similitud de secuencia. Una "ventana de comparación" tal como se utiliza en la presente memoria, se refiere a un segmento de al menos aproximadamente 20 posiciones contiguas, habitualmente de 30 a aproximadamente 75, de 40 a aproximadamente 50, en la que puede compararse una secuencia con una secuencia de referencia del mismo número de posiciones contiguas tras alinear óptimamente las dos secuencias.

Puede realizarse la alineación óptima de secuencias para su comparación utilizando el programa Megalign en la

serie Lasergene de software de bioinformática (DNASTAR, Inc., Madison, WI), utilizando los parámetros por defecto. Este programa realiza varios esquemas de alineación descritos en las siguientes referencias: Dayhoff, M.O. (1978) A model of evolutionary change in proteins - Matrices for detecting distant relationships. En Dayhoff, M.O. (ed.) Atlas of Protein Sequence and Structure, National Biomedical Research Foundation, Washington DC Vol. 5, Sup. 3, págs. 345-358; Hein J. (1990) Unified Approach to Alignment and Phylogenesis págs. 626-645 Methods in Enzymology vol. 183, Academic Press, Inc., San Diego, CA; Higgins, D.G. y Sharp, P.M. (1989) CABIOS 5:151-153; Myers, E.W. y Muller W. (1988) CABIOS 4:11-17; Robinson, E.D. (1971) Comb. Theor 11:105; Santou, N. Nes, M. (1987) Mol. Biol. Evol. 4:406-425; Sneath, P.H.A. y Sokal, R. R. (1973) Numerical Taxonomy - the Principles and Practice of Numerical Taxonomy, Freeman Press, San Francisco, CA; Wilbur, W.J. y Lipman, D.J. (1983) Proc. Natl. Acad., Sci. USA 80:726-730.

Alternativamente, puede realizarse la alineación óptima de secuencias para su comparación mediante el algoritmo de identidad local de Smith y Waterman (1981) Add. APL. Math 2:482, mediante el algoritmo de alineación de identidad de Needleman y Wunsch (1970) J. Mol. Biol. 48:443, mediante los procedimientos de búsqueda de similitud de Pearson y Lipman (1988) Proc. Natl. Acad. Sci. USA 85: 2444, mediante implementaciones informáticas de estos algoritmos (GAP, BESTFIT, BLAST, FASTA, y TFASTA en el paquete de software de Wisconsin Genetics, Genetics Computer Group (GCG), 575 Science Dr., Madison, WI) o mediante inspección.

Un ejemplo preferido de algoritmos que son adecuados para determinar el porcentaje de identidad de secuencia y similitud de secuencia son los algoritmos BLAST y BLAST 2.0, que se describen en Altschul *et al.* (1977) Nucl. Acids Res. 25: 3389-3402 y Altschul *et al.* (1990) J. Mol. Biol. 215:403-410, respectivamente. BLAST y BLAST 2.0 pueden utilizarse, por ejemplo, con los parámetros descritos en la presente memoria, para determinar el porcentaje de identidad de secuencia para los polinucleótidos de la invención. El software para realizar análisis BLAST está disponible al público a través del National Center for Biotechnology Information. En un ejemplo ilustrativo, pueden calcularse puntuaciones acumulativas utilizando, para secuencias de nucleótidos, los parámetros M (puntuación de recompensa para un par de residuos coincidentes; siempre >0) y N (puntuación de penalización para residuos con coincidencia errónea; siempre <0). La extensión de los resultados positivos de palabra se detiene cuando: la puntuación de alineación acumulativa disminuye en una cantidad X desde su valor máximo alcanzado; la puntuación acumulativa llega hasta cero o menos, debido a la acumulación de una o más alineaciones de residuos con puntuación negativa; o se alcanza el final de cualquier secuencia. Los parámetros del algoritmo BLAST, W, T y X, determinan la sensibilidad y velocidad de la alineación. El programa BLASTN (para secuencias de nucleótidos) utiliza como valores por defecto una longitud de palabra (W) de 11, y expectativa (E) de 10, y la matriz de puntuación BLOSUM62 (véase Henikoff y Henikoff (1989) Proc. Natl. Acad. Sci. USA 89:10915) alineaciones, (B) de 50, expectativa (E) de 10, M=5, N=-4 y una comparación de ambas cadenas.

Preferentemente, el "porcentaje de identidad de secuencia" se determina comparando dos secuencias alineadas óptimamente a lo largo de una ventana de comparación de al menos 20 posiciones, en la que la parte de la secuencia de polinucleótido en la ventana de comparación puede comprender adiciones o deleciones (es decir, huecos) del 20 por ciento o menos, habitualmente del 5 al 15 por ciento, o del 10 al 12 por ciento, en comparación con las secuencias de referencia (que no comprenden adiciones o deleciones) para una alineación óptima de las dos secuencias. El porcentaje se calcula determinando el número de posiciones en las que se producen las bases de ácido nucleico idénticas en ambas secuencias para proporcionar el número de posiciones coincidentes, dividiendo el número de posiciones coincidentes entre el número total de posiciones en la secuencia de referencia (es decir, el tamaño de ventana) y multiplicando los resultados por 100 para proporcionar el porcentaje de identidad de secuencia.

Los expertos habituales en la materia apreciarán que, como resultado de la degeneración del código genético, hay muchas secuencias de nucleótidos que codifican para un polipéptido tal como se describe en la presente memoria. Algunos de estos polinucleótidos llevan homología mínima con la secuencia de nucleótidos de cualquier gen nativo. No obstante, la presente invención contempla específicamente polinucleótidos que varían debido a diferencias en la utilización de codones. Además, los alelos de los genes que comprenden las secuencias de polinucleótido proporcionadas en la presente memoria están dentro del alcance de la presente invención. Alelos son genes endógenos que se alteran como resultado de una o más mutaciones, tales como deleciones, adiciones y/o sustituciones de nucleótidos. El ARNm y la proteína resultantes pueden presentar, pero no necesariamente, una estructura o función alteradas. Pueden identificarse alelos utilizando técnicas convencionales (tales como hibridación, amplificación y/o comparación de secuencias de bases de datos).

Por tanto, puede emplearse un enfoque de mutagénesis, tal como mutagénesis específica del sitio, para la preparación de variantes y/o derivados inmunogénicos de los polipéptidos descritos en la presente memoria. Mediante este enfoque, pueden realizarse modificaciones específicas en una secuencia de polipéptido mediante mutagénesis de los polinucleótidos subyacentes que codifican para la misma. Estas técnicas proporcionan un enfoque directo para preparar y someter a prueba variantes de secuencias, por ejemplo, incorporando una o más de las siguientes consideraciones, introduciendo uno o más cambios de secuencia de nucleótidos en el polinucleótido.

La mutagénesis específica del sitio permite la producción de mutantes mediante la utilización de secuencias de oligonucleótido específicas que codifican para la secuencia de ADN de la mutación deseada, así como un número

suficiente de nucleótidos adyacentes, para proporcionar una secuencia de cebador de tamaño y complejidad de secuencia suficientes para formar un dúplex estable en ambos lados de la unión de delección que está atravesándose. Pueden emplearse mutaciones en una secuencia de polinucleótido seleccionada para mejorar, alterar, reducir, modificar o cambiar de otro modo las propiedades del propio polinucleótido, y/o alterar las propiedades, actividad, composición, estabilidad o secuencia primaria del polipéptido codificado.

Los inventores contemplan la mutagénesis de las secuencias de polinucleótido dadas a conocer para alterar una o más propiedades del polipéptido codificado, tales como la inmunogenicidad de una vacuna polipeptídica. Las técnicas de mutagénesis específica del sitio se conocen bien en la materia, y se utilizan ampliamente para crear variantes tanto de polipéptidos como de polinucleótidos. Por ejemplo, con frecuencia se utiliza mutagénesis específica del sitio para alterar una parte específica de una molécula de ADN; se emplea un cebador que normalmente comprende de aproximadamente 14 a aproximadamente 25 nucleótidos aproximadamente de longitud, con de aproximadamente 5 a aproximadamente 10 residuos en ambos lados de la unión de la secuencia que está alterándose.

Tal como apreciarán los expertos en la materia, las técnicas de mutagénesis específica de sitio han empleado con frecuencia un vector de fago que existe en forma tanto monocatenaria como bicatenaria. Los vectores típicos útiles en la mutagénesis dirigida al sitio incluyen vectores tales como el fago M13. Estos fagos están fácilmente disponibles comercialmente y su utilización lo conocen generalmente bien los expertos en la materia. También se emplean de manera rutinaria plásmidos bicatenarios en la mutagénesis dirigida al sitio lo que elimina la etapa de transferir el gen de interés de un plásmido a un fago.

En general, se realiza mutagénesis dirigida al sitio de acuerdo con el presente documento obteniendo en primer lugar un vector monocatenario o separando por fusión dos cadenas de un vector bicatenario que incluye dentro de su secuencia una secuencia de ADN que codifica para el péptido deseado. Se prepara un cebador oligonucleotídico que lleva la secuencia mutada deseada, generalmente de manera sintética. Entonces se aparea este cebador con el vector monocatenario y se somete a enzimas de polimerización de ADN tales como fragmento de Klenow de polimerasa I de *E. coli*, con el fin de completar la síntesis de la cadena que lleva la mutación. Por tanto, se forma un heterodúplex en el que una cadena codifica para la secuencia no mutada original y la segunda cadena lleva la mutación deseada. Entonces se utiliza este vector de heterodúplex para transformar células apropiadas, tales como células de *E. coli*, y se seleccionan clones que incluyen vectores recombinantes que llevan la disposición de secuencia mutada.

La preparación de variantes de secuencia de los segmentos de ADN que codifican para péptido seleccionado utilizando mutagénesis dirigida al sitio proporciona un medio de producción de especies potencialmente útiles y no se pretende que sea limitativo ya que hay otras maneras en las que pueden obtenerse variantes de secuencia de péptidos y las secuencias de ADN que codifican para los mismos. Por ejemplo, pueden tratarse vectores recombinantes que codifican para la secuencia peptídica deseada con agentes mutagénicos, tales como hidroxilamina, para obtener variantes de secuencia. Se encuentran detalles específicos referentes a estos procedimientos y protocolos en las enseñanzas de Maloy *et al.*, 1994; Segal, 1976; Prokop y Bajpai, 1991; Kuby, 1994; y Maniatis *et al.*, 1982.

Tal como se utiliza en la presente memoria, el término "procedimiento de mutagénesis dirigida por oligonucleótido" se refiere a procedimientos dependientes de molde y propagación mediada por vector que dan como resultado un aumento en la concentración de una molécula de ácido nucleico específica con respecto a su concentración inicial, o un aumento en la concentración de una señal detectable, tal como amplificación. Tal como se utiliza en la presente memoria, se pretende que el término "procedimiento de mutagénesis dirigida por oligonucleótido" se refiera a un procedimiento que implica la extensión dependiente de molde de una molécula de cebador. El término procedimiento dependiente de molde se refiere a síntesis de ácido nucleico de una molécula de ARN o de ADN en la que la secuencia de la cadena recién sintetizada de ácido nucleico viene dictada por las reglas bien conocidas de apareamiento de bases complementarias (véase, por ejemplo, Watson, 1987). Normalmente, las metodologías mediadas por vector implican la introducción del fragmento de ácido nucleico en un vector de ADN o ARN, la amplificación clonal del vector y la recuperación del fragmento de ácido nucleico amplificado. Se proporcionan ejemplos de tales metodologías por la patente US n.º 4.237.224.

En otro enfoque para la producción de variantes de polipéptido de la presente invención, puede emplearse recombinación de secuencia recursiva, tal como se describe en la patente US n.º 5.837.458. En este enfoque, se realizan ciclos iterativos de recombinación y examen o selección para "hacer evolucionar" variantes de polinucleótido individuales de la invención que presentan, por ejemplo, actividad inmunogénica potenciada.

En otras realizaciones de la presente invención, las secuencias de polinucleótido proporcionadas en la presente memoria pueden utilizarse ventajosamente como sondas o cebadores para la hibridación de ácido nucleico. Como tales, se contempla que segmentos de ácido nucleico que comprenden o consisten en una región de secuencia de al menos aproximadamente una secuencia contigua de 15 nucleótidos de longitud que presenta la misma secuencia que, o es complementaria a, una secuencia contigua de 15 nucleótidos de longitud dada a conocer en la presente memoria encontrarán utilidad particular. Secuencias complementarias o idénticas contiguas más largas, por ejemplo,

## ES 2 405 405 T3

aquellas de aproximadamente 20, 30, 40, 50, 100, 200, 500, 1000 (incluyendo todas las longitudes intermedias) e incluso hasta secuencias de longitud completa también serán de utilidad en determinadas realizaciones.

5 La capacidad de tales sondas de ácido nucleico para hibridarse específicamente con una secuencia de interés les permitirá ser útiles en la detección de la presencia de secuencias complementarias en una muestra dada. Sin embargo, también se consideran otros usos, tales como la utilización de la información de secuencia para la preparación de cebadores de especies mutantes, o cebadores para su utilización en la preparación de otros constructos genéticos.

10 Moléculas de polinucleótido que presentan regiones de secuencia que consisten en tramos de nucleótidos contiguos de 10-14, 15-20, 30, 50 o incluso de 100-200 nucleótidos aproximadamente (incluyendo también longitudes intermedias), idénticos o complementarios a una secuencia de polinucleótido dada a conocer en la presente memoria, se contemplan particularmente como sondas de hibridación para su utilización, por ejemplo, en  
15 en transferencia de tipo Southern y Northern. Esto permitirá analizar un producto génico, o fragmento del mismo, tanto en diversos tipos de células como también en diversas células bacterianas. El tamaño total del fragmento, así como el tamaño del/de los tramo(s) complementario(s), dependerá en última instancia de la utilización o la aplicación pretendida del segmento de ácido nucleico particular. Fragmentos más pequeños encontrarán utilización generalmente en realizaciones de hibridación, en las que la longitud de la región complementaria contigua puede variarse, tal como entre aproximadamente 15 y aproximadamente 100 nucleótidos, pero pueden utilizarse tramos de  
20 complementariedad contiguos más grandes, según la longitud de secuencias complementarias que se desee detectar.

25 La utilización de una sonda de hibridación de aproximadamente 15-25 nucleótidos de longitud permite la formación de una molécula dúplex que es tanto estable como selectiva. Sin embargo, se prefieren generalmente moléculas que presentan secuencias complementarias contiguas a lo largo de tramos de más de 15 bases de longitud, con el fin de aumentar la estabilidad y selectividad del híbrido, y así mejorar la calidad y el grado de moléculas híbridas específicas obtenidas. Generalmente se preferirá diseñar moléculas de ácido nucleico que presentan tramos complementarios génicos de 15 a 25 nucleótidos contiguos o incluso más largos cuando se desee.

30 Pueden seleccionarse sondas de hibridación de cualquier parte de cualquiera de las secuencias dadas a conocer en la presente memoria. Lo único que se necesita es revisar las secuencias expuestas en la presente memoria, o cualquier parte contigua de las secuencias, de desde aproximadamente 15-25 nucleótidos de longitud hasta, e incluyendo, la secuencia de longitud completa, que se desee utilizar como sonda o cebador. La elección de secuencias de sonda y cebador puede estar regida por diversos factores. Por ejemplo, puede desearse emplear  
35 cebadores hacia los extremos terminales de la secuencia total.

Pueden prepararse fácilmente fragmentos o segmentos de polinucleótido pequeños, por ejemplo, sintetizando directamente el fragmento mediante medios químicos, tal como se pone comúnmente en práctica utilizando un sintetizador de oligonucleótidos automático. Además, pueden obtenerse fragmentos mediante aplicación de  
40 tecnología de reproducción de ácido nucleico, tal como la tecnología de PCR™ de la patente US 4.683.202, introduciendo secuencias seleccionadas en vectores recombinantes para la producción recombinante, y mediante otras técnicas de ADN recombinante generalmente conocidas por los expertos en la materia de biología molecular.

45 Las secuencias de nucleótidos de la invención pueden utilizarse por su capacidad para formar selectivamente moléculas dúplex con tramos complementarios de todo el gen o fragmentos de gen de interés. Dependiendo de la aplicación considerada, normalmente se deseará emplear condiciones de hibridación variables para lograr grados variables de selectividad de sonda hacia la secuencia diana. Para aplicaciones que requieren alta selectividad, normalmente se deseará emplear condiciones relativamente rigurosas para formar los híbridos, por ejemplo, se seleccionarán condiciones de contenido en sal relativamente bajo y/o temperatura alta, tal como se proporcionan por  
50 una concentración de sal de desde aproximadamente 0,02 M hasta aproximadamente 0,15 M de sal a temperaturas de desde aproximadamente 50°C hasta aproximadamente 70°C. Tales condiciones selectivas toleran poca, si es que toleran, coincidencia errónea entre la sonda y la cadena de molde o diana, y serán particularmente adecuadas para aislar secuencias relacionadas.

55 Evidentemente, para algunas aplicaciones, por ejemplo, en las que se desea preparar mutantes empleando una cadena de cebador de mutante hibridada con un molde subyacente, se necesitarán normalmente condiciones de hibridación menos rigurosas (rigurosidad reducida) con el fin de permitir la formación del heterodúplex. En estas circunstancias, puede desearse emplear condiciones de sal tales como las de desde aproximadamente 0,15 M hasta aproximadamente 0,9 M de sal, a temperaturas que oscilan entre aproximadamente 20°C y aproximadamente 55°C.  
60 De ese modo pueden identificarse fácilmente especies con hibridación cruzada como señales de hibridación positivas con respecto a hibridaciones de control. En cualquier caso, generalmente se aprecia que las condiciones pueden volverse más rigurosas mediante la adición de cantidades crecientes de formamida, que sirve para desestabilizar el dúplex híbrido de la misma manera que el aumento de temperatura. Por tanto, pueden manipularse fácilmente las condiciones de hibridación y, por tanto, generalmente serán un procedimiento de elección dependiendo de los resultados deseados.  
65

Según otra realización de la presente invención, se proporcionan composiciones de polinucleótido que comprenden oligonucleótidos antisentido. Se ha demostrado que los oligonucleótidos antisentido son inhibidores eficaces y dirigidos de síntesis de proteína, y, por consiguiente, proporcionan un enfoque terapéutico mediante el cual puede tratarse una enfermedad inhibiendo la síntesis de proteínas que contribuyen a la enfermedad. La eficacia de los oligonucleótidos antisentido para inhibir la síntesis de proteínas está bien establecida. Por ejemplo, la síntesis de poligalacturonasa y el receptor de acetilcolina muscarínico tipo 2 se inhiben mediante oligonucleótidos antisentido dirigidos a sus respectivas secuencias de ARNm (patente US 5.739.119 y patente US 5.759.829). Además, se han demostrado ejemplos de inhibición antisentido con la proteína nuclear ciclina, el gen resistencia a múltiples fármacos (MDG1), 1CAM-1, E-selectina, STK-1, receptor de GABAA estriatal y EGF humano (Jaskulski *et al.*, Science. 10 de junio de 1988; 240(4858):1544-6; Vasanthakumar y Ahmed, Cancer Commun. 1989; 1(4): 225-32; Peris *et al.*, Brain Res Mol Brain Res. 15 de junio de 1998; 57(2):310-20; patente US 5.801.154; patente US 5.789.573; patente US 5.718.709 y patente US 5.610.288). También se han descrito constructos antisentido que inhiben y pueden utilizarse para tratar una variedad de proliferaciones celulares anómalas, por ejemplo cáncer (patente US 5.747.470; patente US 5.591.317 y patente US 5.783.683).

Por tanto, en la presente memoria se dan a conocer secuencias de oligonucleótido que comprenden la totalidad, o una parte, de cualquier secuencia que pueda unirse específicamente a una secuencia de polinucleótido descrita en la presente memoria, o un complemento de la misma. Los oligonucleótidos antisentido pueden comprender ADN o derivados del mismo. Los oligonucleótidos pueden comprender ARN o derivados del mismo. Los oligonucleótidos pueden ser ADN modificados que comprenden una estructura principal modificada con fosforotioato. Las secuencias de oligonucleótido pueden comprender ácidos nucleicos peptídicos o derivados de los mismos. En cada caso, las composiciones preferidas comprenden una región de secuencia que es complementaria, y más preferentemente sustancialmente complementaria, y de manera incluso más preferible completamente complementaria a una o más partes de polinucleótidos dados a conocer en la presente memoria. La selección de composiciones antisentido específicas para una secuencia génica dada se basa en el análisis de la secuencia diana elegida y la determinación de la estructura secundaria,  $T_m$ , energía de unión y estabilidad relativa. Pueden seleccionarse composiciones antisentido basándose en su incapacidad relativa para formar dímeros, horquillas u otras estructuras secundarias que reducirían o impedirían la unión específica al ARNm diana en una célula huésped. Regiones diana altamente preferidas del ARNm son aquellas que están en el, o cerca del, codón de iniciación de la traducción AUG, y aquellas secuencias que son sustancialmente complementarias a regiones 5' del ARNm. Estos análisis de estructura secundaria y consideraciones de selección de sitio diana pueden realizarse, por ejemplo, utilizando la v. 4 del software de análisis de cebadores OLIGO y/o el software de algoritmo BLASTN 2.0.5 (Altschul *et al.*, Nucleic Acids Res. 1997, 25(17):3389-402).

También se contempla la utilización de un procedimiento de administración antisentido que emplea un vector de péptido corto, denominado MPG (27 residuos). El péptido MPG contiene un dominio hidrófobo derivado de la secuencia de fusión de gp41 del VIH y un dominio hidrófilo de la secuencia de localización nuclear del antígeno T de SV40 (Morris *et al.*, Nucleic Acids Res. 15 de julio de 1997; 25(14):2730-6). Se ha demostrado que varias moléculas del péptido MPG recubren los oligonucleótidos antisentido y pueden administrarse en células de mamífero en cultivo en menos de 1 hora con una eficacia relativamente alta (90%). Además, la interacción con MPG aumenta fuertemente tanto la estabilidad del oligonucleótido frente a nucleasa como la capacidad de atravesar la membrana plasmática.

Las composiciones de polinucleótido descritas en la presente memoria pueden utilizarse en el diseño y la preparación de moléculas de ribozima para inhibir la expresión de los polipéptidos y proteínas bacterianos de la presente invención en células bacterianas. Las ribozimas son complejos de ARN-proteína que escinden ácidos nucleicos de una manera específica del sitio. Las ribozimas presentan dominios catalíticos específicos que presentan actividad endonucleasa (Kim y Cech, Proc Natl Acad Sci USA. diciembre de 1987; 84(24): 8788-92; Forster y Symons, Cell. 24 de abril de 1987; 49(2):211-20). Por ejemplo, un gran número de ribozimas aceleran reacciones de transferencia de fosfoéster con un alto grado de especificidad, con frecuencia escindiendo sólo uno de varios fosfoésteres en un sustrato de oligonucleótido (Cech *et al.*, Cell. diciembre de 1981; 27(3 Pt 2):487-96; Michel y Westhof, J Mol Biol. 5 de diciembre de 1990; 216(3):585-610; Reinhold-Hurek y Shub, Nature. 14 de mayo de 1992; 357(6374):173-6). Esta especificidad se ha atribuido al requisito de que el sustrato se una mediante interacciones de apareamiento de bases específico a la secuencia de guía interna ("IGS") de la ribozima antes de la reacción química.

En la actualidad se conocen seis variedades básicas de ARN enzimáticos que se producen de manera natural. Cada uno puede catalizar la hidrólisis de enlaces fosfodiéster de ARN en trans (y por tanto pueden escindir otras moléculas de ARN) en condiciones fisiológicas. En general, los ácidos nucleicos enzimáticos actúan en primer lugar uniéndose a un ARN diana. Tal unión se produce a través de la parte de unión a diana de un ácido nucleico enzimático que se mantiene en estrecha proximidad con una parte enzimática de la molécula que actúa para escindir el ARN diana. Por tanto, en primer lugar el ácido nucleico enzimático reconoce y después se une a un ARN diana mediante apareamiento de bases complementarias, y una vez unido al sitio correcto, actúa enzimáticamente para cortar el ARN diana. La escisión estratégica de un ARN diana de este tipo destruirá su capacidad para dirigir la síntesis de una proteína codificada. Tras haberse unido un ácido nucleico enzimático y haber escindido su diana de ARN, se libera del ARN para buscar otra diana y puede unirse y escindir repetidamente nuevas dianas.

La naturaleza enzimática de una ribozima es ventajosa con respecto a muchas tecnologías, tales como tecnología antisentido (en la que una molécula de ácido nucleico simplemente se une a una diana de ácido nucleico para bloquear su traducción) ya que la concentración de ribozima necesaria para realizar un tratamiento terapéutico es inferior a la de un oligonucleótido antisentido. Esta ventaja refleja la capacidad de la ribozima para actuar enzimáticamente. Por tanto, una única molécula de ribozima puede escindir muchas moléculas de ARN diana. Además, la ribozima es un inhibidor altamente específico, dependiendo la especificidad de inhibición no sólo del mecanismo de apareamiento de bases de unión al ARN diana, sino también del mecanismo de escisión de ARN diana. Apareamientos erróneos individuales, o sustituciones de bases, cerca del sitio de escisión pueden eliminar completamente la actividad catalítica de una ribozima. Apareamientos erróneos similares en moléculas antisentido no impiden su acción (Woolf *et al.*, Proc Natl Acad Sci USA. 15 de agosto de 1992; 89(16):7305-9). Por tanto, la especificidad de acción de una ribozima es superior a la de un oligonucleótido antisentido que se une al mismo sitio de ARN.

La molécula de ácido nucleico enzimático puede formarse en un motivo de cabeza de martillo, horquilla, virus de la hepatitis  $\delta$ , intrón de grupo I o ARN de ARNasa P (en asociación con una secuencia guía de ARN) o ARN de VS de *Neurospora*. Se describen ejemplos de motivos de cabeza de martillo por Rossi *et al.* Nucleic Acids Res. 11 de septiembre de 1992; 20(17):4559-65. Se describen ejemplos de motivos de horquilla por Hampel *et al.* (publicación de solicitud de patente europea n.º EP 0360257), Hampel y Tritz, Biochemistry 13 de junio de 1989; 28(12):4929-33; Hampel *et al.*, Nucleic Acids Res. 25 de enero de 1990; 18(2):299-304 y la patente US 5.631.359. Se describe un ejemplo del motivo de virus de la hepatitis  $\delta$  por Perrotta y Been, Biochemistry. 1 de diciembre de 1992; 31(47):11843-52; se describe un ejemplo del motivo de ARNasa P por Guerrier-Takada *et al.*, Cell. diciembre de 1983; 35(3 Pt 2):849-57; se describe un motivo de ribozima de ARN de VS de *Neurospora* por Collins (Saville y Collins, Cell. 18 de mayo de 1990; 61(4):685-96; Saville y Collins, Proc Natl Acad Sci USA. 1 de octubre de 1991; 88(19):8826-30; Collins y Olive, Biochemistry. 23 de marzo de 1993; 32(11):2795-9); y se describe un ejemplo del intrón de grupo I en (patente US 4.987.071). Lo único que importa en una molécula de ácido nucleico enzimático de esta invención es que presente un sitio de unión a sustrato específico que sea complementario a una o más de las regiones de ARN del gen diana, y que presente secuencias de nucleótidos dentro o que rodean al sitio de unión a sustrato que confieran una actividad de escisión de ARN a la molécula. Por tanto no es necesario limitar los constructos de ribozima a motivos específicos mencionados en la presente memoria.

Pueden diseñarse ribozimas tal como se describe en la publicación de solicitud de patente internacional n.º WO 93/23569 y la publicación de solicitud de patente internacional n.º WO 94/02595 y sintetizarse para someterse a pruebas *in vitro* e *in vivo*, tal como se describe. También pueden optimizarse tales ribozimas para su administración. Aunque se proporcionan ejemplos específicos, los expertos en la materia reconocerán que pueden utilizarse dianas de ARN equivalentes de otras especies cuando sea necesario.

Puede optimizarse la actividad de ribozima alterando la longitud de las ramas de unión de ribozima, o sintetizando químicamente ribozimas con modificaciones que impiden su degradación por ribonucleasas séricas (véase por ejemplo, la publicación de solicitud de patente internacional n.º WO 92/07065; la publicación de solicitud de patente internacional n.º WO 93/15187; la publicación de solicitud de patente internacional n.º WO 91/03162; la publicación de solicitud de patente europea n.º 92110298.4; la patente US 5.334.711 y la publicación de solicitud de patente internacional n.º WO 94/13688, que describen diversas modificaciones químicas que pueden realizarse en restos de azúcar de moléculas de ARN enzimáticas), modificaciones que potencian su eficacia en células, y eliminación de bases del tallo II para acortar los tiempos de síntesis de ARN y reducir los requisitos químicos.

Sullivan *et al.* (publicación de solicitud de patente internacional n.º WO 94/02595) describen los procedimientos generales para administrar moléculas de ARN enzimático. Pueden administrarse ribozimas a células mediante una variedad de procedimientos conocidos por los expertos en la materia, incluyendo, pero sin limitarse a, encapsulación en liposomas, mediante iontoforesis o mediante incorporación en otros vehículos, tales como hidrogeles, ciclodextrinas, nanocápsulas biodegradables y microesferas bioadhesivas. Para algunas indicaciones, pueden administrarse ribozimas directamente *ex vivo* a células o tejidos con o sin los vehículos mencionados anteriormente. Alternativamente, puede administrarse localmente la combinación de ARN/vehículo mediante inhalación directa, mediante inyección directa o mediante la utilización de un catéter, bomba de infusión o endoprótesis. Otras vías de administración incluyen, pero no se limitan a, inyección intravascular, intramuscular, subcutánea o en articulaciones, inhalación de aerosol, administración oral (forma de comprimido o pastilla), tópica, sistémica, ocular, intraperitoneal y/o intratecal. Se proporcionan descripciones más detalladas de suministro y administración de ribozimas en la publicación de solicitud de patente internacional n.º WO 94/02595 y la publicación de solicitud de patente internacional n.º WO 93/23569.

Otro medio de acumulación de altas concentraciones de una(s) ribozima(s) dentro de células es incorporar las secuencias que codifican para ribozima en un vector de expresión de ADN. La transcripción de las secuencias de ribozima se impulsan a partir de un promotor para ARN polimerasa I (pol I), ARN polimerasa II (pol II) o ARN polimerasa III (pol III) de eucariota. Los transcritos de promotores de pol II o pol III se expresarán a altos niveles en todas las células; los niveles de un promotor de pol II dado en un tipo de célula dado dependerán de la naturaleza de las secuencias reguladoras génicas (potenciadores, silenciadores, etc.) presentes en la proximidad. También

pueden utilizarse promotores de ARN polimerasa de procarionta, siempre que la enzima ARN polimerasa de procarionta se exprese en las células apropiadas. Se ha mostrado que ribozimas expresadas a partir de tales promotores funcionan en células de mamífero. Tales unidades de transcripción pueden incorporarse en una variedad de vectores para la introducción en células de mamífero, incluyendo, pero sin limitarse a, vectores de ADN de plásmido, vectores de ADN viral (tales como vectores de adenovirus o adenoasociados) o vectores de ARN viral (tales como vectores retrovirales, de virus del bosque Semliki, de virus Sindbis).

En otra realización de la invención, se proporcionan composiciones de ácidos nucleicos peptídicos (ANP). ANP es un mimético de ADN en el que las nucleobases están unidas a una estructura principal pseudopeptídica (Good y Nielsen, *Antisense Nucleic Acid Drug Dev.* 1997 7(4) 431-37). Puede utilizarse ANP en varios procedimientos que tradicionalmente han utilizado ARN o ADN. Con frecuencia, las secuencias de ANP funcionan mejor en técnicas que las correspondientes secuencias de ARN o de ADN y presentan utilidades que no son inherentes a ARN o ADN. Se proporciona una revisión de ANP incluyendo procedimientos de preparación, características de, y procedimientos de utilización, por Corey (*Trends Biotechnol* junio de 1997; 15(6):224-9). Como tal, en determinadas realizaciones, pueden prepararse secuencias de ANP que son complementarias a una o más partes de la secuencia de ARNm de ACE, y tales composiciones de ANP pueden utilizarse para regular, alterar, disminuir o reducir la traducción de ARNm específico de ACE, y así alterar el nivel de actividad ACE en una célula huésped a la que se le han administrado tales composiciones de ANP.

Los ANP presentan enlaces 2-aminoetil-glicina que sustituyen a la estructura principal de fosfodiéster normal del ADN (Nielsen *et al.*, *Science* 6 de diciembre de 1991; 254(5037):1497-500; Hanvey *et al.*, *Science*. 27 de noviembre de 1992; 258(5087)1 481-5; Hyrup y Nielsen, *Bioorg Med Chem.* enero de 1996; 4(1):5-23). Esta química presenta tres consecuencias importantes: en primer lugar, al contrario que el ADN u oligonucleótidos de fosforotioato, los ANP son moléculas neutras; en segundo lugar, los ANP son acirales, lo que evita la necesidad de desarrollar una síntesis estereoselectiva; y en tercer lugar, la síntesis de ANP utiliza protocolos de Boc o Fmoc convencionales para la síntesis de péptidos en fase sólida, aunque se han utilizado otros procedimientos, incluyendo un procedimiento de Merrifield modificado.

Hay monómeros de ANP u oligómeros ya preparados disponibles comercialmente de PerSeptive Biosystems (Framingham, MA). La síntesis de ANP mediante protocolos o bien de Boc o bien de Fmoc es directa utilizando protocolos manuales o automáticos (Norton *et al.*, *Bioorg Med Chem.* abril de 1995; 3(4):437-45). El protocolo manual se presta a la producción de ANP químicamente modificados o a la síntesis simultánea de familias de ANP estrechamente relacionadas.

Como con la síntesis de péptidos, el éxito de una síntesis de ANP particular dependerá de las propiedades de la secuencia elegida. Por ejemplo, aunque en teoría los ANP pueden incorporar cualquier combinación de bases de nucleótidos, la presencia de purinas adyacentes puede conducir a deleciones de uno o más residuos en el producto. Dado que se espera esta dificultad, se sugiere que, en la producción de ANP con purinas adyacentes, debe repetirse el acoplamiento de residuos que es probable que se añadan de manera ineficaz. Esto debe ir seguido por la purificación de ANP mediante cromatografía de líquidos a alta presión de fase inversa, proporcionando rendimientos y pureza de producto similares a los observados durante la síntesis de péptidos.

Pueden lograrse modificaciones de ANP para una aplicación dada mediante el acoplamiento de aminoácidos durante la síntesis en fase sólida o mediante la unión de compuestos que contienen un grupo ácido carboxílico a la amina N-terminal expuesta. Alternativamente, pueden modificarse ANP tras la síntesis mediante acoplamiento con una lisina o cisteína introducida. La facilidad con la que pueden modificarse los ANP facilita la optimización para una mejor solubilidad o para requisitos funcionales específicos. Una vez sintetizados, puede confirmarse la identidad de ANP y sus derivados mediante espectrometría de masas. Diversos estudios han realizado y utilizado modificaciones de ANP (por ejemplo, Norton *et al.*, *Bioorg Med Chem.* abril de 1995; 3(4):437-45; Petersen *et al.*, *J Pept Sci.* mayo-junio de 1995; 1(3):175-83; Orum *et al.*, *Biotechniques.* septiembre de 1995; 19(3):472-80; Footer *et al.*, *Biochemistry.* 20 de agosto de 1996; 35(33); 10673-9; Griffith *et al.*, *Nucleic Acids Res.* 11 de agosto de 1995; 23(15):3003-8; Pardridge *et al.*, *Proc Natl Acad Sci USA.* 6 de junio de 1995; 92(12):5592-6; Boffa *et al.*, *Proc Natl Acad Sci USA.* 14 de marzo de 1995; 92(6); 1901-5; Gambacorti-Passerini *et al.*, *Blood.* 15 de agosto de 1996; 88(4):1411-7; Armitage *et al.*, *Proc Natl Acad Sci USA.* 11 de noviembre de 1997; 94(23); 12320-5; Seeger *et al.*, *Biotechniques.* septiembre de 1997; 23(3):512-7). La patente US n.º 5.700.922 se refiere a moléculas quiméricas de ANP-ADN-ANP y a sus utilidades en diagnóstico, modulación de proteínas en organismos y tratamiento de estados sensibles a agentes terapéuticos.

Se comentan procedimientos de caracterización de propiedades de unión antisentido de ANP en Rose (*Anal Chem.* 15 de diciembre de 1993; 65(24):3545-9) y Jensen *et al.* (*Biochemistry.* 22 de abril de 1997; 36(16):5072-7). Rose utiliza electroforesis en gel capilar para determinar la unión de ANP a su oligonucleótido complementario, midiendo la estequiometría y la cinética de unión relativa. Se realizaron tipos de mediciones similares por Jensen *et al.* utilizando tecnología BIAcore™.

Otras aplicaciones de ANP que se han descrito y que resultarán evidentes para el experto en la materia incluyen la utilización en invasión de ADN, inhibición antisentido, análisis mutacional, potenciadores de la transcripción,

purificación de ácido nucleico, aislamiento de genes transcripcionalmente activos, bloqueo de la unión de factor de transcripción, escisión de genoma, biosensores, hibridación *in situ*, y similares.

#### Identificación, caracterización y expresión de polinucleótidos

5 Pueden identificarse, prepararse y/o manipularse composiciones de polinucleótidos de la presente invención utilizando cualquiera de una variedad de técnicas bien establecidas (véase generalmente, Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1989, y otras referencias similares). Por ejemplo, puede identificarse un polinucleótido, tal como se describe con más detalle a  
10 continuación, examinando una micromatriz de ADNc para detectar ADNc bacterianos presentes en muestras de tejido aisladas de individuos afectados por EII en comparación con muestras aisladas de individuos no afectados. Tales exámenes pueden realizarse, por ejemplo, utilizando la tecnología de micromatriz de Affymetrix, Inc. (Santa Clara, CA) según las instrucciones del fabricante (y esencialmente tal como se describe por Schena *et al.*, Proc. Natl. Acad. Sci. USA 93:10614-10619, 1996 y Heller *et al.*, Proc. Natl. Acad. Sci. USA 94:2150-2155, 1997).  
15 Alternativamente, pueden identificarse composiciones de polinucleótido de la presente invención examinando bibliotecas de expresión de cizalladura al azar genómica de bacterias cecales de ratón o de ser humano, tal como se describe en el ejemplo 1.

20 Alternativamente, pueden amplificarse polinucleótidos a partir de ADNc preparado a partir de células que expresan las proteínas bacterianas descritas en la presente memoria. Muchos procedimientos dependientes de molde están disponibles para amplificar una secuencia diana de interés presente en una muestra. Uno de los procedimientos de amplificación mejor conocidos es la reacción en cadena de la polimerasa (PCR<sup>TM</sup>) que se describe en detalle en las patentes US n.ºs 4.683.195, 4.683.202 y 4.800.159. En resumen, en la PCRT<sup>TM</sup>, se preparan dos secuencias de cebadores que son complementarias a regiones en cadenas complementarias opuestas de la secuencia diana. Se  
25 añade un exceso de desoxinucleósido trifosfatos a una mezcla de reacción junto con una ADN polimerasa (por ejemplo, polimerasa de Taq). Si la secuencia diana está presente en una muestra, los cebadores se unirán a la diana y la polimerasa hará que los cebadores se extiendan a lo largo de la secuencia diana añadiendo nucleótidos. Aumentando o reduciendo la temperatura de la mezcla de reacción, los cebadores extendidos se disociarán de la diana para formar productos de reacción, los cebadores en exceso se unirán a la diana y al producto de reacción y se repite el procedimiento. Preferentemente puede realizarse un procedimiento de amplificación mediante PCR<sup>TM</sup> y transcripción inversa con el fin de cuantificar la cantidad de ARNm amplificado. En la materia se conocen bien las metodologías de reacción en cadena de la polimerasa.

35 Cualquiera de varios otros procedimientos dependientes de molde, muchos de los cuales son variaciones de la técnica de amplificación mediante PCR<sup>TM</sup>, se conocen fácilmente y están disponibles en la materia. De manera ilustrativa, algunos de tales procedimientos incluyen la reacción en cadena de la ligasa (denominada LCR), descrita, por ejemplo, en la publicación de solicitud de patente europea n.º 320.308 y la patente US n.º 4.883.750; replicasa Qbeta, descrita en la publicación de solicitud de patente internacional PCT n.º PCT/US87/00880; amplificación por desplazamiento de cadena (SDA) y reacción de reparación de cadena (RCR). Todavía otros procedimientos de  
40 amplificación se describen en la solicitud de patente de Gran Bretaña n.º 2 202 328 y en la publicación de solicitud de patente internacional PCT n.º PCT/US89/01025. Otros procedimientos de amplificación de ácido nucleico incluyen sistemas de amplificación basados en la transcripción (TAS) (publicación de solicitud de patente internacional PCT n.º WO 88/10315), incluyendo amplificación basada en secuencia de ácido nucleico (NASBA) y 3SR. La publicación de solicitud de patente europea n.º 329.822 describe un procedimiento de amplificación de ácido nucleico que implica sintetizar de manera cíclica ARN monocatenario ("ARNmc"), ADNmc y ADN bicatenario (ADNbc). La publicación de solicitud de patente internacional PCT n.º WO 89/06700 describe un esquema de amplificación de secuencia de ácido nucleico basado en la hibridación de una secuencia de promotor/cebador con un ADN monocatenario ("ADNmc") diana seguido por la transcripción de muchas copias de ARN de la secuencia. Otros procedimientos de amplificación tales como "RACE" (Frohman, 1990) y "PCR unilateral" (Ohara, 1989)  
50 también los conocen bien los expertos en la materia.

Puede utilizarse una parte amplificada de un polinucleótido de la presente invención para aislar un gen de longitud completa a partir de una biblioteca adecuada (por ejemplo, una biblioteca de ADNc bacteriano) utilizando técnicas bien conocidas. Dentro de tales técnicas, se examina una biblioteca (ADNc o genómica) utilizando una o más  
55 sondas o cebadores polinucleotídicos adecuados para la amplificación. Preferentemente, se selecciona el tamaño de una biblioteca para incluir moléculas más grandes. También pueden preferirse bibliotecas con cebadores al azar para identificar regiones en sentido 5' y anteriores de genes. Se prefieren bibliotecas genómicas para obtener intrones y secuencias en sentido 5' en extensión.

60 Para las técnicas de hibridación, puede marcarse una secuencia parcial (por ejemplo, mediante traslado de mellas o marcaje en el extremo con <sup>32</sup>P) utilizando técnicas bien conocidas. Entonces se examina generalmente una biblioteca bacteriana o de bacteriófagos hibridando filtros que contienen colonias bacterianas desnaturalizadas (o céspedes que contienen placas de lisis de fagos) con la sonda marcada (véase Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1989). Se seleccionan placas de lisis o colonias hibridantes y se expanden, y se aísla el ADN para su análisis adicional. Pueden analizarse clones de  
65 ADNc para determinar la cantidad de secuencia adicional, por ejemplo, mediante PCR utilizando un cebador de la

secuencia parcial y un cebador del vector. Pueden generarse mapas de restricción y secuencias parciales para identificar uno o más clones solapantes. Entonces puede determinarse la secuencia completa utilizando técnicas convencionales, que pueden implicar generar una serie de clones de delección. Entonces pueden ensamblarse las secuencias solapantes resultantes en una única secuencia contigua. Puede generarse una molécula de ADNc de longitud completa ligando fragmentos adecuados, utilizando técnicas bien conocidas.

Alternativamente, técnicas de amplificación, tales como las descritas anteriormente, pueden ser útiles para obtener una secuencia codificante de longitud completa a partir de una secuencia de ADNc parcial. Una técnica de amplificación de este tipo es la PCR inversa (véase Triglia *et al.*, Nucl. Acids Res. 16:8186, 1988), que utiliza enzimas de restricción para generar un fragmento en la región conocida del gen. Entonces se forma un círculo con el fragmento mediante ligación intramolecular y se utiliza como molde para PCR con cebadores divergentes derivados de la región conocida. Dentro de un enfoque alternativo, pueden recuperarse secuencias adyacentes a una secuencia parcial mediante amplificación con un cebador para una secuencia de ligador y un cebador específico para una región conocida. Normalmente se someten las secuencias amplificadas a un segundo ciclo de amplificación con el mismo cebador de ligador y un segundo cebador específico para la región conocida. Una variación de este procedimiento, que emplea dos cebadores que inician la extensión en sentidos opuestos desde la secuencia conocida, se describe en el documento WO 96/38591. Otra técnica de este tipo se conoce como "amplificación rápida de extremos de ADNc" o RACE. Esta técnica implica la utilización de un cebador interno y un cebador externo, que se hibrida con una región de poliA o secuencia de vector, para identificar secuencias que son 5' y 3' con respecto a una secuencia conocida. Las técnicas adicionales incluyen PCR de captura (Lagerstrom *et al.*, PCR Methods Applic. 1:111-19, 1991) y PCR de paseo (Parker *et al.*, Nucl. Acids. Res. 19:3055-60, 1991). También pueden emplearse otros procedimientos que emplean amplificación para obtener una secuencia de ADNc de longitud completa.

En determinados casos, es posible obtener una secuencia de ADNc de longitud completa mediante análisis de secuencias proporcionadas en una base de datos de etiqueta de secuencia expresada (EST), tal como la disponible de GenBank. Generalmente pueden realizarse búsquedas de EST solapantes utilizando programas bien conocidos (por ejemplo, búsquedas con BLAST de NCBI), y tales EST pueden utilizarse para generar una secuencia de longitud completa contigua. También pueden obtenerse secuencias de ADN de longitud completa mediante análisis de fragmentos genómicos.

En otras realizaciones de la invención, pueden utilizarse secuencias de polinucleótido o fragmentos las mismas que codifican para polipéptidos de la invención, o proteínas de fusión o equivalentes funcionales de las mismas, en moléculas de ADN recombinante para dirigir la expresión de un polipéptido en células huésped apropiadas. Debido a la degeneración inherente del código genético, pueden producirse otras secuencias de ADN que codifican sustancialmente para la misma secuencia de aminoácidos o para una funcionalmente equivalente y pueden utilizarse estas secuencias para clonar y expresar un polipéptido dado.

Tal como entenderán los expertos en la materia, puede ser ventajoso en algunos casos producir secuencias de nucleótidos que codifican para polipéptidos que presentan codones que no se producen de manera natural. Por ejemplo, pueden seleccionarse codones preferidos por un huésped procarionta o eucariota particular para aumentar la tasa de expresión de proteína o para producir un transcrito de ARN recombinante que presenta propiedades deseables, tales como una semivida que es mayor que la de un transcrito generado a partir de la secuencia que se produce de manera natural.

Además, las secuencias de polinucleótido de la presente invención pueden diseñarse mediante ingeniería utilizando procedimientos generalmente conocidos en la materia con el fin de alterar secuencias que codifican para polipéptidos por una variedad de motivos, incluyendo, pero sin limitarse a, alteraciones que modifican la clonación, el procesamiento y/o la expresión del producto génico. Por ejemplo, puede utilizarse intercambio de ADN mediante fragmentación al azar y reensamblaje mediante PCR de fragmentos génicos y oligonucleótidos sintéticos para diseñar mediante ingeniería las secuencias de nucleótidos. Además, puede utilizarse mutagénesis dirigida al sitio para insertar nuevos sitios de restricción, alterar patrones de glicosilación, cambiar preferencia de codones, producir variantes de corte y empalme o introducir mutaciones, etcétera.

En otra realización de la invención, pueden ligarse secuencias de ácido nucleico naturales, modificadas o recombinantes a una secuencia heteróloga para codificar para una proteína de fusión. Por ejemplo, para examinar bibliotecas de péptidos para detectar inhibidores de actividad de polipéptido, puede ser útil codificar para una proteína química que pueda reconocerse por un anticuerpo disponible comercialmente. También puede diseñarse una proteína de fusión mediante ingeniería para contener un sitio de escisión ubicado entre la secuencia que codifica para polipéptido y la secuencia de proteína heteróloga, de modo que puede escindirse el polipéptido y separarse mediante purificación del resto heterólogo.

Pueden sintetizarse secuencias que codifican para un polipéptido deseado, en su totalidad o en parte, utilizando procedimientos químicos bien conocidos en la materia (véase 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). Alternativamente, la propia proteína puede producirse utilizando procedimientos químicos para sintetizar la secuencia de aminoácidos de un polipéptido,

o una parte de la misma. Por ejemplo, puede realizarse síntesis de péptidos utilizando diversas técnicas en fase sólida (Roberge, J. Y. *et al.* (1995) *Science* 269:202-204) y puede lograrse la síntesis automática, por ejemplo, utilizando el sintetizador de péptidos ABI 431A (Perkin Elmer, Palo Alto, CA).

5 Un péptido recién sintetizado puede purificarse sustancialmente mediante cromatografía de líquidos de alto rendimiento preparativa (por ejemplo, Creighton, T. (1983) *Proteins, Structures and Molecular Principles*, WH Freeman and Co., Nueva York, N.Y.) u otras técnicas comparables disponibles en la materia. La composición de los péptidos sintéticos puede confirmarse mediante secuenciación o análisis de aminoácidos (por ejemplo, el procedimiento de degradación de Edman). Además, la secuencia de aminoácidos de un polipéptido, o cualquier parte de la misma, puede alterarse durante síntesis directa y/o combinarse utilizando procedimientos químicos con secuencias de otras proteínas, o cualquier parte de las mismas, para producir un polipéptido variante.

15 Con el fin de expresar un polipéptido deseado, pueden insertarse las secuencias de nucleótidos que codifican para el polipéptido, 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 codificante insertada. Pueden utilizarse procedimientos que conocen bien los expertos en la materia para construir vectores de expresión que contienen secuencias que codifican para un polipéptido de interés y elementos de control de la transcripción y la traducción apropiados. Estos procedimientos incluyen técnicas de ADN recombinante *in vitro*, técnicas sintéticas, y recombinación genética *in vivo*. Tales técnicas se describen, por ejemplo, en Sambrook, J. *et al.* (1989) *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Press, Plainview, N.Y., y Ausubel, F. M. *et al.* (1989) *Current Protocols in Molecular Biology*, John Wiley & Sons, Nueva York, N.Y.

25 Puede utilizarse una variedad de sistemas de vector de expresión/huésped para contener y expresar secuencias de polinucleótido. Estos incluyen, pero no se limitan a, microorganismos tales como bacterias transformadas con vectores de expresión de bacteriófago recombinante, plásmido o ADN de cósmido; levaduras transformadas con vectores de expresión de levadura; sistemas de células de insecto infectadas con vectores de expresión virales (por ejemplo, baculovirus); sistemas de células vegetales transformadas con vectores de expresión virales (por ejemplo, virus del mosaico de la coliflor, VMC; virus del mosaico del tabaco, VMT) o con vectores de expresión bacterianos (por ejemplo, plásmidos Ti o pBR322); o sistemas de células animales.

30 Los "elementos de control" o "secuencias reguladoras" presentes en un vector de expresión son aquellas regiones no traducidas del vector (potenciadores, promotores, regiones no traducidas en sentido 5' y 3') que interactúan con proteínas de células huésped para llevar a cabo la transcripción y traducción. Tales elementos pueden variar en cuanto a su intensidad y especificidad. Dependiendo del sistema de vector y del huésped utilizados, pueden utilizarse varios elementos de transcripción y traducción adecuados, incluyendo promotores constitutivos e inducibles. Por ejemplo, cuando se clonan en sistemas bacterianos, pueden utilizarse promotores inducibles tales como el promotor de lacZ híbrido del fagémido pBLUESCRIPT (Stratagene, La Jolla, Calif.) o plásmido pSPORT1 (Gibco BRL, Gaithersburg, MD) y similares. En sistemas de células de mamíferos, se prefieren generalmente promotores de genes de mamíferos o de virus de mamíferos. Si es necesario generar una línea celular que contenga múltiples copias de la secuencia que codifica para un polipéptido, pueden utilizarse ventajosamente vectores basados en SV40 o VEB con un marcador seleccionable apropiado.

45 En sistemas bacterianos, puede seleccionarse cualquiera de varios vectores de expresión dependiendo de la utilización pretendida para el polipéptido expresado. Por ejemplo, cuando se necesitan grandes cantidades, por ejemplo para la inducción de anticuerpos, pueden utilizarse vectores que dirigen la expresión a alto nivel de proteínas de fusión que se purifican fácilmente. Tales vectores incluyen, pero no se limitan a, los vectores de clonación y expresión de *E. coli* multifuncionales tales como pBLUESCRIPT (Stratagene), en el que la secuencia que codifica para el polipéptido de interés puede ligarse en el vector en marco con secuencias para la Met amino-terminal y los 7 residuos posteriores de beta-galactosidasa de modo que se produce una proteína híbrida; vectores pIN (Van Heeke, G. y S. M. Schuster (1989) *J. Biol. Chem.* 264:5503-5509); y similares. También pueden utilizarse vectores pGEX (Promega, Madison, Wis.) para expresar polipéptidos foráneos como proteínas de fusión con glutatión S-transferasa (GST). En general, tales proteínas de fusión son solubles y pueden purificarse fácilmente a partir de células lisadas mediante adsorción a perlas de glutatión-agarosa seguido por la elución en presencia de glutatión libre. Pueden diseñarse proteínas preparadas en tales sistemas para incluir sitios de escisión de heparina, trombina o factor XA proteasa de modo que el polipéptido clonado de interés puede liberarse del resto de GST a voluntad.

60 En la levadura, *Saccharomyces cerevisiae*, pueden utilizarse varios vectores que contienen promotores constitutivos o inducibles tales como factor alfa, alcohol oxidasa y PGH. Para revisiones, véanse Ausubel *et al.* (citado anteriormente) y Grant *et al.* (1987) *Methods Enzymol.* 153:516-544.

65 En casos en los que se utilizan vectores de expresión de plantas, la expresión de secuencias que codifican para polipéptidos puede impulsarse mediante cualquiera de varios promotores. Por ejemplo, pueden utilizarse promotores virales tales como los promotores 35S y 19S de VMC solos o en combinación con la secuencia líder omega de VMT (Takamatsu, N. (1987) *EMBO J.* 6: 307-311). Alternativamente, pueden utilizarse promotores de plantas tales como la subunidad pequeña de RUBISCO o promotores de choque térmico (Coruzzi, G. *et al.* (1984) *EMBO J.* 3:1671-

1680; Broglie, R. *et al.* (1984) *Science* 224:838-843; y Winter, J. *et al.* (1991) *Results Probl. Cell Differ.* 17:85-105). Estos constructos pueden introducirse en células vegetales mediante transformación de ADN directa o transfección mediada por patógeno. Tales técnicas se describen en varias revisiones generalmente disponibles (véase, por ejemplo, Hobbs, S. o Murry, L. E. en McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, Nueva York, N.Y.; págs. 191-196).

También puede utilizarse un sistema de insecto para expresar un polipéptido de interés. Por ejemplo, en un sistema de este tipo, se utiliza virus de la poliedrosis nuclear de *Autographa californica* (AcNPV) como vector para expresar genes foráneos en células de *Spodoptera frugiperda* o en larvas de *Trichoplusia*. Las secuencias que codifican para el polipéptido pueden clonarse en una región no esencial del virus, tal como el gen de poliedrina, y colocarse bajo el control del promotor de poliedrina. La inserción satisfactoria de la secuencia que codifica para polipéptido hará que el gen de la poliedrina se vuelva inactivo y producirá virus recombinante que carece de proteína de recubrimiento. Entonces pueden utilizarse los virus recombinantes para infectar, por ejemplo, células de *S. frugiperda* o larvas de *Trichoplusia* en las que puede expresarse el polipéptido de interés (Engelhard, E. K. *et al.* (1994) *Proc. Natl. Acad. Sci.* 91:3224-3227).

En células huésped de mamífero, están generalmente disponibles varios sistemas de expresión basados en virus. Por ejemplo, en casos en los que se utiliza un adenovirus como vector de expresión, pueden ligarse secuencias que codifican para un polipéptido de interés en un complejo de transcripción/traducción de adenovirus que consiste en el promotor tardío y la secuencia líder tripartito. La inserción en una región no esencial E1 o E3 del genoma viral puede utilizarse para obtener un virus viable que puede expresar el polipéptido en células huésped infectadas (Logan, J. y Shenk, T. (1984) *Proc. Natl. Acad. Sci.* 81:3655-3659). Además, pueden utilizarse potenciadores de la transcripción, tales como el potenciador del virus de sarcoma de Rous (VSR), para aumentar la expresión en células huésped de mamífero.

También pueden utilizarse señales de iniciación específicas para lograr una traducción más eficaz de secuencias que codifican para un polipéptido de interés. Tales señales incluyen el codón de iniciación ATG y secuencias adyacentes. En casos en los que se insertan secuencias que codifican para el polipéptido, su codón de iniciación y secuencias en sentido 5' en el vector de expresión apropiado, pueden no necesitarse señales de control de la transcripción o la traducción adicionales. Sin embargo, en casos en los que sólo se inserta la secuencia codificante, o una parte de la misma, deben proporcionarse señales de control de la traducción exógenas que incluyen el codón de iniciación ATG. Además, el codón de iniciación debe estar en el marco de lectura correcto para garantizar la traducción de todo el inserto. Los codones de iniciación y elementos de traducción exógenos pueden ser de diversos orígenes, tanto naturales como sintéticos. La eficacia de la expresión puede potenciarse mediante la inclusión de potenciadores que son apropiados para el sistema celular particular que se utiliza, tal como los descritos en la bibliografía (Scharf, D. *et al.* (1994) *Results Probl. Cell Differ.* 20:125-162).

Además, puede elegirse una cepa de célula huésped por su capacidad para modular la expresión de las secuencias insertadas o para procesar la proteína expresada de la manera deseada. Tales modificaciones del polipéptido incluyen, pero no se limitan a, acetilación, carboxilación, glicosilación, fosforilación, lipidación y acilación. También puede utilizarse procesamiento postraduccional que escinde una forma "prepro" de la proteína para facilitar la correcta inserción, plegamiento y/o función. Pueden elegirse diferentes células huésped tales como CHO, COS, HeLa, MDCK, HEK293 y WI38, que presentan mecanismos característicos y maquinaria celular específica para tales actividades postraducionales, para garantizar la correcta modificación y procesamiento de la proteína foránea.

Para la producción de alto rendimiento, a largo plazo, de proteínas recombinantes, generalmente se prefiere la expresión estable. Por ejemplo, pueden transformarse líneas celulares que expresan de manera estable un polinucleótido de interés utilizando vectores de expresión que pueden contener orígenes de replicación virales y/o elementos de expresión endógenos y un gen marcador seleccionable en el mismo vector o en uno separado. Tras la introducción del vector, pueden dejarse crecer las células durante 1-2 días en unos medios enriquecidos antes de cambiarse a medios selectivos. El propósito del marcador seleccionable es conferir resistencia a la selección, y su presencia permite el crecimiento y la recuperación de células que expresan satisfactoriamente las secuencias introducidas. Los clones resistentes de células transformadas de manera estable pueden proliferar utilizando técnicas de cultivo tisular apropiadas para el tipo de célula.

Puede utilizarse cualquiera de varios sistemas de selección para recuperar líneas celulares transformadas. Estos incluyen, pero no se limitan a, los genes de timidina cinasa del virus del herpes simple (Wigler, M. *et al.* (1977) *Cell* 11:223-32) y adenina fosforibosiltransferasa (Lowy, I. *et al.* (1990) *Cell* 22:817-23) que pueden emplearse en células tk.sup.- o aprt.sup.-, respectivamente. Además, puede utilizarse resistencia a antimetabolitos, antibióticos o herbicidas como base para la selección; por ejemplo, dhfr que confiere resistencia a metotrexato (Wigler, M. *et al.* (1980) *Proc. Natl. Acad. Sci.* 77:3567-70); npt, que confiere resistencia a los aminoglucósidos, neomicina y G-418 (Colbere-Garapin, F. *et al.* (1981) *J. Mol. Biol.* 150: 1-14); y als o pat, que confieren resistencia a clorsulfurón y fosfotricina acetiltransferasa, respectivamente (Murry, citado anteriormente). Se han descrito genes seleccionables adicionales, por ejemplo, trpB, que permite a células utilizar indol en lugar de triptófano, o hisD, que permite a células utilizar histinol en lugar de histidina (Hartman, S. C. y R. C. Mulligan (1988) *Proc. Natl. Acad. Sci.* 85:8047-51). La utilización de marcadores visibles ha ganado popularidad utilizándose ampliamente marcadores tales como

antocianinas, beta-glucuronidasa y su sustrato GUS, y luciferasa y su sustrato luciferina, no sólo para identificar transformantes, sino también para cuantificar la cantidad de expresión de proteína transitoria o estable que puede atribuirse a un sistema de vector específico (Rhodes, C. A. *et al.* (1995) *Methods Mol. Biol.* 55:121-131).

5 Aunque la presencia/ausencia de expresión de un gen marcador sugiere que el gen de interés también está presente, puede necesitar confirmarse su presencia y expresión. Por ejemplo, si se inserta la secuencia que codifica para un polipéptido dentro de una secuencia de gen marcador, pueden identificarse células recombinantes que contienen secuencias mediante la ausencia de función de gen marcador. Alternativamente, puede colocarse un gen  
10 marcador en tándem con una secuencia que codifica para polipéptido bajo el control de un único promotor. La expresión del gen marcador en respuesta a la inducción o selección también indica habitualmente la expresión del gen en tándem.

15 Alternativamente, pueden identificarse células huésped que contienen y expresan una secuencia de polinucleótido deseada mediante una variedad de procedimientos conocidos por los expertos en la materia. Estos procedimientos incluyen, pero no se limitan a, hibridaciones ADN-ADN o ADN-ARN y bioensayo de proteínas o técnicas de inmunoensayo que incluyen, por ejemplo, tecnologías basadas en membrana, disolución o chip para la detección y/o cuantificación de ácido nucleico o proteína.

20 En la materia se conoce una variedad de protocolos para detectar y medir la expresión de productos codificados por polinucleótido, utilizando anticuerpos o bien policlonales o bien monoclonales específicos para el producto. Los ejemplos incluyen ensayo de inmunoabsorción ligado a enzimas (ELISA), radioinmunoensayo (RIA) y clasificación celular activada por fluorescencia (FACS). Para algunas aplicaciones puede preferirse un inmunoensayo basado en anticuerpos monoclonales, de dos sitios, utilizando anticuerpos monoclonales reactivos frente a dos epítopos no  
25 interferentes en un polipéptido dado, pero también puede emplearse un ensayo de unión competitiva. Estos y otros ensayos se describen, entre otros lugares, en Hampton, R. *et al.* (1990; *Serological Methods, a Laboratory Manual*, APS Press, St Paul, Minn.) y Maddox, D. E. *et al.* (1983; *J. Exp. Med.* 158:1211-1216).

30 Los expertos en la materia conocen una amplia variedad de marcadores y técnicas de conjugación y pueden utilizarse en diversos ensayos de ácidos nucleicos y aminoácidos. Los medios para producir sondas de PCR o hibridación marcadas para detectar secuencias relacionadas con polinucleótidos incluyen oligomarcaje, traslado de mellas, marcaje de extremos o amplificación por PCR utilizando un nucleótido marcado. Alternativamente, las secuencias, o cualquier parte de las mismas, pueden clonarse en un vector para la producción de una sonda de ARNm. Tales vectores se conocen en la materia, están disponibles comercialmente, y pueden utilizarse para sintetizar sondas de ARN *in vitro* mediante adición de una ARN polimerasa apropiada tal como T7, T3 o SP6 y  
35 nucleótidos marcados. Estos procedimientos pueden realizarse utilizando una variedad de kits disponibles comercialmente. Los marcadores o moléculas indicadoras adecuadas, que pueden utilizarse incluyen radionúclidos, enzimas, agentes fluorescentes, quimioluminiscentes o cromogénicos así como sustratos, cofactores, inhibidores, partículas magnéticas, y similares.

40 Pueden cultivarse células huésped transformadas con una secuencia de polinucleótido de interés en condiciones adecuadas para la expresión y la recuperación de la proteína a partir del cultivo celular. La proteína producida por una célula recombinante puede secretarse o contenerse intracelularmente dependiendo de la secuencia y/o el vector utilizados. Tal como entenderán los expertos en la materia, pueden diseñarse vectores de expresión que contienen polinucleótidos de la invención para contener secuencias señal que dirigen la secreción del polipéptido codificado a  
45 través de una membrana de célula procarionota o eucariota. Pueden utilizarse otros constructos recombinantes para unir secuencias que codifican para un polipéptido de interés a una secuencia de nucleótidos que codifica para un dominio de polipéptido que facilitará la purificación de proteínas solubles. Tales dominios que facilitan la purificación incluyen, pero no se limitan a, péptidos quelantes de metales tales como módulos de histidina-triptófano que permiten la purificación sobre metales inmovilizados, dominios de proteína A que permiten la purificación sobre inmunoglobulina inmovilizada, y el dominio utilizado en el sistema de purificación por extensión/afinidad de FLAGS (Immunex Corp., Seattle, Wash.). Puede utilizarse la inclusión de secuencias de ligador escindibles tales como las especificadas para factor XA o enterocinasa (Invitrogen, San Diego, Calif.) entre el dominio de purificación y el polipéptido codificado para facilitar la purificación. Un vector de expresión de este tipo proporciona la expresión de  
50 una proteína de fusión que contiene un polipéptido de interés y un ácido nucleico que codifica para 6 residuos de histidina que preceden a un sitio de escisión de tioredoxina o enterocinasa. Los residuos de histidina facilitan la purificación en IMIAC (cromatografía por afinidad de iones metálicos inmovilizados) tal como se describe en Porath, J. *et al.* (1992, *Prot. Exp. Purif.* 3:263-281) mientras que el sitio de escisión de enterocinasa proporciona un medio para purificar el polipéptido deseado a partir de la proteína de fusión. Se proporciona una evaluación de vectores que contienen proteínas de fusión en Kroll, D. J. *et al.* (1993; *DNA Cell Biol.* 12:441-453).

60 Además de procedimientos de producción recombinantes, pueden producirse polipéptidos de la invención, y fragmentos de los mismos, mediante síntesis de péptidos directa utilizando técnicas en fase sólida (Merrifield J. (1963) *J. Am. Chem. Soc.* 85:2149-2154). Puede realizarse la síntesis de proteínas utilizando técnicas manuales o mediante automatización. Puede lograrse la síntesis automática, por ejemplo, utilizando el sintetizador de péptidos Applied Biosystems 431A (Perkin Elmer). Alternativamente, pueden sintetizarse químicamente diversos fragmentos por separado y combinarse utilizando procedimientos químicos para producir la molécula de longitud completa.

Composiciones de anticuerpos, fragmentos de los mismos y otros agentes de unión

Según otro aspecto, la presente invención proporciona además un anticuerpo aislado, o fragmento de unión a antígeno del mismo, que se une específicamente al polipéptido de SEC ID N°: 79. También se describen en la presente memoria agentes de unión, tales como anticuerpos y fragmentos de unión a antígeno de los mismos, que muestran unión inmunológica a un polipéptido bacteriano dado a conocer en la presente memoria, o a una parte, variante o derivado del mismo. Algunos anticuerpos dados a conocer en la presente memoria se unen a un receptor tipo Toll. Los receptores tipo Toll (TLR) ilustrativos incluyen, pero no se limitan a, TLR5. Los anticuerpos de la presente invención pueden unirse a una proteína de flagelina. En otra realización, los anticuerpos de la presente invención son anticuerpos neutralizantes que bloquean la interacción entre TLR5 y una proteína de flagelina.

Se dice que un anticuerpo, o fragmento de unión a antígeno del mismo, se “une específicamente a”, “se une inmunológicamente a” y/o “es inmunológicamente reactivo frente a” un polipéptido de la invención si reacciona a un nivel detectable (dentro de, por ejemplo, un ensayo ELISA) con el polipéptido, y no reacciona de manera detectable con polipéptidos no relacionados en condiciones similares.

La unión inmunológica, tal como se utiliza en este contexto, se refiere de manera general a las interacciones no covalentes del tipo que se producen entre una molécula de inmunoglobulina y un antígeno para el cual la inmunoglobulina es específica. La resistencia o afinidad de interacciones de unión inmunológica puede expresarse en cuanto a la constante de disociación ( $K_d$ ) de la interacción, en la que una  $K_d$  menor representa una mayor afinidad. Las propiedades de unión inmunológica de polipéptidos seleccionados pueden cuantificarse utilizando procedimientos bien conocidos en la materia. Un procedimiento de este tipo conlleva medir las velocidades de formación y disociación de complejo sitio de unión a antígeno/antígeno, en el que esas velocidades dependen de las concentraciones de las parejas del complejo, de la afinidad de la interacción y de parámetros geométricos que influyen por igual sobre la velocidad en ambos sentidos. Por tanto, pueden determinarse tanto la “constante de velocidad de formación” ( $K_{on}$ ) como la “constante de velocidad de disociación” ( $K_{off}$ ) mediante cálculo de las concentraciones y las velocidades reales de asociación y disociación. La razón de  $K_{off}/K_{on}$  permite cancelar todos los parámetros no relacionados con la afinidad, y por tanto es igual a la constante de disociación  $K_d$ . Véase, generalmente, Davies *et al.* (1990) Annual Rev. Biochem. 59:439-473.

Un “sitio de unión a antígeno”, o “parte de unión”, de un anticuerpo se refiere a la parte de la molécula de inmunoglobulina que participa en la unión al antígeno. El sitio de unión a antígeno está formado por residuos de aminoácido de las regiones variables N-terminales (“V”) de las cadenas pesada (“H”) y ligera (“L”). Tres tramos altamente divergentes dentro de las regiones V de las cadenas pesada y ligera se denominan “regiones hipervariables” que están interpuestas entre tramos flanqueantes más conservados conocidos como “regiones de entramado”, o “FR”. Por tanto el término “FR” se refiere a secuencias de aminoácidos que se encuentran de manera natural entre, y adyacentes a, regiones hipervariables en inmunoglobulinas. En una molécula de anticuerpo, las tres regiones hipervariables de una cadena ligera y las tres regiones hipervariables de una cadena pesada están dispuestas unas con respecto a otras en un espacio tridimensional para formar una superficie de unión a antígeno. La superficie de unión a antígeno es complementaria a la superficie tridimensional de un antígeno unido, y las tres regiones hipervariables de cada una de las cadenas pesada y ligera se denominan “regiones determinantes de complementariedad”, o “CDR”.

Además, puede que los agentes de unión puedan diferenciar entre pacientes con y sin EII, utilizando los ensayos representativos proporcionados en la presente memoria. Por ejemplo, anticuerpos u otros agentes de unión que se unen a una proteína bacteriana generarán preferentemente una señal que indica la presencia de EII en al menos aproximadamente el 20% de los pacientes con la enfermedad, más preferentemente al menos aproximadamente el 30% de los pacientes. Alternativamente, o además, el anticuerpo generará una señal negativa que indica la ausencia de la enfermedad en al menos aproximadamente el 90% de los individuos sin EII. Para determinar si un agente de unión satisface este requisito, pueden someterse a ensayo muestras biológicas (por ejemplo, sangre, sueros, esputo, orina, heces y/o biopsias) de pacientes con y sin EII (según se determina utilizando pruebas clínicas convencionales) tal como se describe en la presente memoria para detectar la presencia de polipéptidos que se unen al agente de unión. Preferentemente, se someterá a ensayo un número estadísticamente significativo de muestras con y sin la enfermedad. Cada agente de unión debe satisfacer los criterios anteriores; sin embargo, los expertos habituales en la materia reconocerán que pueden utilizarse agentes de unión en combinación para mejorar la sensibilidad.

Además, puede que los agentes de unión puedan identificar a pacientes en riesgo de desarrollar EII, utilizando los ensayos representativos proporcionados en la presente memoria. Por ejemplo, anticuerpos u otros agentes de unión que se unen a una proteína bacteriana generarán preferentemente una señal que indica un riesgo de desarrollar EII en al menos aproximadamente el 20% de los pacientes con historia familiar positiva de la enfermedad, o pacientes con un riesgo genético definido de desarrollar la enfermedad (por ejemplo, individuos positivos para la mutación NOD2), más preferentemente al menos aproximadamente el 30% de dichos individuos. Alternativamente, o además, el anticuerpo generará una señal negativa que indica la ausencia de riesgo de desarrollar la enfermedad en al menos aproximadamente el 90% de los individuos sin historia familiar o sin riesgo genético definido de desarrollar

- 5 EII. Para determinar si un agente de unión satisface este requisito, pueden someterse a ensayo muestras biológicas (por ejemplo, sangre, sueros, esputo, orina, heces y/o biopsias) de pacientes con y sin factores de riesgo para el desarrollo de EII (según se determina utilizando pruebas clínicas convencionales) tal como se describe en la presente memoria para detectar la presencia de polipéptidos que se unen al agente de unión. Preferentemente, se someterá a ensayo un número estadísticamente significativo de muestras de individuos con y sin riesgo de desarrollar la enfermedad. Cada agente de unión debe satisfacer los criterios anteriores; sin embargo, los expertos habituales en la materia reconocerán que pueden utilizarse agentes de unión en combinación para mejorar la sensibilidad.
- 10 Pueden utilizarse agentes de unión de la presente invención adicionalmente, o bien solos o bien en combinación con otras modalidades de diagnóstico, para subdividir a pacientes con EII en categorías de enfermedad que serán sensibles o resistentes a tratamientos nuevos o existentes.
- 15 Cualquier agente que satisfaga los requisitos anteriores puede ser un agente de unión. Por ejemplo, un agente de unión puede ser un ribosoma, con o sin un componente peptídico, una molécula de ARN o un polipéptido. Preferentemente, un agente de unión es un anticuerpo o un fragmento de unión a antígeno del mismo. Pueden prepararse anticuerpos mediante una variedad de técnicas conocidas por los expertos habituales en la materia. Véase, por ejemplo, Harlow y Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. En general, pueden producirse anticuerpos mediante técnicas de cultivo celular, incluyendo la generación de anticuerpos monoclonales tal como se describe en la presente memoria, o mediante transfección de genes de anticuerpos en huéspedes de células bacterianas o de mamífero adecuadas, con el fin de permitir la producción de anticuerpos recombinantes. En una técnica, se inyecta inicialmente un inmunógeno que comprende el polipéptido en cualquiera de una amplia variedad de mamíferos (por ejemplo, ratones, ratas, conejos, ovejas o cabras). En esta etapa, los polipéptidos de esta invención pueden servir como inmunógeno sin modificación. Alternativamente, particularmente para polipéptidos relativamente cortos, puede provocarse una respuesta inmunitaria superior si se une el polipéptido a una proteína portadora, tal como albúmina sérica bovina o hemocianina de lapa californiana. Se inyecta el inmunógeno en el huésped animal, preferentemente según un calendario predeterminado que incorpora una o más inmunizaciones de refuerzo, y se extrae sangre periódicamente de los animales. Entonces pueden purificarse anticuerpos policlonales específicos para el polipéptido a partir de tales antisueros, por ejemplo, mediante cromatografía por afinidad utilizando el polipéptido acoplado a un soporte sólido adecuado.
- 20 Pueden prepararse anticuerpos monoclonales específicos para un polipéptido antigénico de interés, por ejemplo, utilizando la técnica de Kohler y Milstein, *Eur. J. Immunol.* 6:511-519, 1976, y mejoras de la misma. En resumen, estos procedimientos implican la preparación de líneas celulares inmortales que pueden producir anticuerpos que presentan la especificidad deseada (es decir, reactividad con el polipéptido de interés). Tales líneas celulares pueden producirse, por ejemplo, a partir de esplenocitos obtenidos de un animal inmunizado tal como se describió anteriormente. Entonces se immortalizan los esplenocitos, por ejemplo, mediante fusión con una pareja de fusión de célula de mieloma, preferentemente una que es singénica con el animal inmunizado. Puede emplearse una variedad de técnicas de fusión. Por ejemplo, pueden combinarse los esplenocitos y las células de mieloma con un detergente no iónico durante algunos minutos y después sembrarse en placas a baja densidad sobre un medio selectivo que soporta el crecimiento de células híbridas, pero no de células de mieloma. Una técnica de selección preferida utiliza selección con HAT (hipoxantina, aminopterina, timidina). Tras un tiempo suficiente, habitualmente de aproximadamente 1 a 2 semanas, se observan colonias de híbridos. Se seleccionan colonias individuales y se someten a una prueba sus sobrenadantes de cultivo para determinar actividad de unión frente al polipéptido. Se prefieren hibridomas que presentan alta reactividad y especificidad.
- 35 Pueden aislarse anticuerpos monoclonales a partir de los sobrenadantes de colonias de hibridoma en crecimiento. Además, pueden emplearse diversas técnicas para potenciar el rendimiento, tales como inyección de la línea celular de hibridoma en la cavidad peritoneal de un huésped vertebrado adecuado, tal como un ratón. Entonces pueden recogerse anticuerpos monoclonales del líquido ascítico o de la sangre. Pueden eliminarse contaminantes de los anticuerpos mediante técnicas convencionales, tales como cromatografía, filtración en gel, precipitación y extracción. Los polipéptidos de esta invención pueden utilizarse en el procedimiento de purificación, por ejemplo, en una etapa de cromatografía por afinidad.
- 40 Pueden prepararse anticuerpos monoclonales específicos para un polipéptido antigénico de interés, por ejemplo, utilizando la técnica de Kohler y Milstein, *Eur. J. Immunol.* 6:511-519, 1976, y mejoras de la misma. En resumen, estos procedimientos implican la preparación de líneas celulares inmortales que pueden producir anticuerpos que presentan la especificidad deseada (es decir, reactividad con el polipéptido de interés). Tales líneas celulares pueden producirse, por ejemplo, a partir de esplenocitos obtenidos de un animal inmunizado tal como se describió anteriormente. Entonces se immortalizan los esplenocitos, por ejemplo, mediante fusión con una pareja de fusión de célula de mieloma, preferentemente una que es singénica con el animal inmunizado. Puede emplearse una variedad de técnicas de fusión. Por ejemplo, pueden combinarse los esplenocitos y las células de mieloma con un detergente no iónico durante algunos minutos y después sembrarse en placas a baja densidad sobre un medio selectivo que soporta el crecimiento de células híbridas, pero no de células de mieloma. Una técnica de selección preferida utiliza selección con HAT (hipoxantina, aminopterina, timidina). Tras un tiempo suficiente, habitualmente de aproximadamente 1 a 2 semanas, se observan colonias de híbridos. Se seleccionan colonias individuales y se someten a una prueba sus sobrenadantes de cultivo para determinar actividad de unión frente al polipéptido. Se prefieren hibridomas que presentan alta reactividad y especificidad.
- 45 Pueden aislarse anticuerpos monoclonales a partir de los sobrenadantes de colonias de hibridoma en crecimiento. Además, pueden emplearse diversas técnicas para potenciar el rendimiento, tales como inyección de la línea celular de hibridoma en la cavidad peritoneal de un huésped vertebrado adecuado, tal como un ratón. Entonces pueden recogerse anticuerpos monoclonales del líquido ascítico o de la sangre. Pueden eliminarse contaminantes de los anticuerpos mediante técnicas convencionales, tales como cromatografía, filtración en gel, precipitación y extracción. Los polipéptidos de esta invención pueden utilizarse en el procedimiento de purificación, por ejemplo, en una etapa de cromatografía por afinidad.
- 50 En la materia se conocen varias moléculas terapéuticamente útiles que comprenden sitios de unión a antígeno que pueden mostrar propiedades de unión inmunológica de una molécula de anticuerpo. La enzima proteolítica papaína escinde de manera preferencial moléculas de IgG para producir varios fragmentos, dos de los cuales (los fragmentos "F(ab)") comprenden cada uno un heterodímero covalente que incluye un sitio de unión a antígeno intacto. La enzima pepsina puede escindir moléculas de IgG para proporcionar varios fragmentos, incluyendo el fragmento "F(ab)<sup>1</sup>"<sub>2</sub> que comprende ambos sitios de unión a antígeno. Puede producirse un fragmento "Fv" mediante escisión proteolítica preferencial de una molécula de inmunoglobulina IgM, y en ocasiones poco frecuentes IgG o IgA. Sin embargo, los fragmentos Fv se derivan más comúnmente utilizando técnicas recombinantes conocidas en la materia. El fragmento Fv incluye un heterodímero V<sub>H</sub>::V<sub>L</sub> no covalente que incluye un sitio de unión a antígeno que conserva gran parte de las capacidades de reconocimiento y unión a antígeno de la molécula de anticuerpo nativa.
- 55 Inbar *et al.* (1972) *Proc. Nat. Acad. Sci. USA* 69: 2659-2662; Hochman *et al.* (1976) *Biochem* 15:2706-2710; y Ehrlich *et al.* (1980) *Biochem* 19:4091-4096.

Un polipéptido de Fv de cadena sencilla ("sFv") es un heterodímero  $V_H::V_L$  covalentemente unido que se expresa a partir de una fusión génica que incluye genes que codifican para  $V_H$  y para  $V_L$  unidos mediante un ligador que codifica para péptido. Huston *et al.* (1988) Proc. Nat. Acad. Sci. USA 85(16):5879-5883. Se han descrito varios procedimientos para distinguir estructuras químicas para convertir las cadenas de polipéptido ligera y pesada agregadas de manera natural (pero separadas químicamente) de una región V de anticuerpo en una molécula sFv que se plegará en una estructura tridimensional sustancialmente similar a la estructura de un sitio de unión a antígeno. Véanse, por ejemplo, las patentes US n.ºs 5.091.513 y 5.132.405, concedidas a Huston *et al.*; y la patente US n.º 4.946.778, concedida a Ladner *et al.*

Cada una de las moléculas descritas anteriormente incluye un conjunto de CDR de cadena pesada y de cadena ligera, respectivamente interpuestos entre un conjunto de FR de cadena pesada y de cadena ligera FR que proporciona soporte para las CDR y que define la relación espacial de las CDR unas con respecto a otras. Tal como se utiliza en la presente memoria, el término "conjunto de CDR" se refiere a las tres regiones hipervariables de una región V de cadena pesada o ligera. Avanzando desde el extremo N-terminal de una cadena pesada o ligera, estas regiones se denominan "CDR1", "CDR2" y "CDR3" respectivamente. Por tanto, un sitio de unión a antígeno incluye seis CDR, que comprenden el conjunto de CDR de cada una de la región V de cadena pesada y de ligera. Un polipéptido que comprende una única CDR, (por ejemplo, una CDR1, CDR2 o CDR3) se denomina en la presente memoria "unidad de reconocimiento molecular". Análisis cristalográficos de varios complejos antígeno-anticuerpo han demostrado que los residuos de aminoácido de CDR forman un amplio contacto con antígeno unido, en los que el contacto con antígeno más amplio es con la CDR3 de cadena pesada. Por tanto, las unidades de reconocimiento molecular son principalmente responsables de la especificidad de un sitio de unión a antígeno.

Tal como se utiliza en la presente memoria, el término "conjunto de FR" se refiere a las cuatro secuencias de aminoácidos flanqueantes que enmarcan las CDR de un conjunto de CDR de una región V de cadena pesada o ligera. Algunos residuos de FR pueden entrar en contacto con antígeno unido; sin embargo, las FR son principalmente responsables del plegamiento de la región V en el sitio de unión a antígeno, particularmente los residuos de FR directamente adyacentes a las CDR. Dentro de las FR, determinados residuos de amino y determinadas características estructurales se conservan muy altamente. Con respecto a esto, todas las secuencias de la región V contienen un bucle de disulfuro interno de aproximadamente 90 residuos de aminoácido. Cuando las regiones V se pliegan en un sitio de unión, las CDR se presentan como motivos de bucle salientes que forman una superficie de unión a antígeno. Generalmente se reconoce que hay regiones estructurales conservadas de FR que influyen en la forma plegada de los bucles de CDR en determinadas estructuras "canónicas" (independientemente de la secuencia de aminoácidos de CDR precisa). Además, se sabe que determinados residuos de FR participan en contactos entre dominios no covalentes que estabilizan la interacción de las cadenas pesadas y ligeras de anticuerpos.

Se han descrito varias moléculas de anticuerpo "humanizadas" que comprenden un sitio de unión a antígeno derivado de una inmunoglobulina no humana, incluyendo anticuerpos quiméricos que presentan regiones V de roedor y sus CDR asociadas fusionadas con dominios constantes humanos (Winter *et al.* (1991) Nature 349:293-299; Lobuglio *et al.* (1989) Proc. Nat. Acad. Sci. USA 86:4220-4224; Shaw *et al.* (1987) J Immunol. 138:4534-4538; y Brown *et al.* (1987) Cancer Res. 47:3577-3583), CDR de roedor injertadas en una FR de soporte humana antes de la fusión con un dominio constante de anticuerpo humano apropiado (Riechmann *et al.* (1988) Nature 332:323-327; Verhoeyen *et al.* (1988) Science 239:1534-1536; y Jones *et al.* (1986) Nature 321:522-525) y CDR de roedor soportadas por FR de roedor recombinantemente recubiertas (publicación de patente europea n.º 519.596, publicada el 23 de diciembre de 1992). Estas moléculas "humanizadas" se diseñan para minimizar la respuesta inmunológica no deseada frente a moléculas de anticuerpo de roedor anti-ser humano que limita la duración y eficacia de aplicaciones terapéuticas de estos restos en receptores humanos.

Tal como se utiliza en la presente memoria, los términos "FR recubiertas" y "FR recombinantemente recubiertas" se refieren a la sustitución selectiva de residuos de FR de, por ejemplo, una región V de cadena pesada o ligera de roedor, por residuos de FR humanos con el fin de proporcionar una molécula xenogénica que comprende un sitio de unión a antígeno que conserva sustancialmente la totalidad de la estructura de plegamiento de polipéptido de FR nativa. Las técnicas de recubrimiento se basan en el entendimiento de que las características de unión a ligando de un sitio de unión a antígeno están determinadas principalmente por la estructura y la disposición relativa de los conjuntos de CDR de cadena pesada y ligera dentro de la superficie de unión a antígeno. Davies *et al.* (1990) Ann. Rev. Biochem. 59:439-473. Por tanto, sólo puede conservarse la especificidad de unión a antígeno en un anticuerpo humanizado en el que se mantiene cuidadosamente las estructuras de CDR, su interacción entre sí y su interacción con el resto de los dominios de región V. Utilizando técnicas de recubrimiento, se sustituyen selectivamente residuos de FR exteriores (por ejemplo, accesibles a disolvente) que se encuentran fácilmente por el sistema inmunitario por residuos humanos para proporcionar una molécula híbrida que comprende una superficie recubierta débilmente inmunogénica o sustancialmente no inmunogénica.

El procedimiento de recubrimiento hace uso de los datos de secuencia disponibles para dominios variables de anticuerpos humanos compilados por Kabat *et al.*, en Sequences of Proteins of Immunological Interest, 4ª ed., (U.S. Dept. of Health and Human Services, U.S. Government Printing Office, 1987), actualizaciones de la base de datos

de Kabat, y otras bases de datos accesibles estadounidenses y extranjeras (tanto de ácido nucleico como de proteína). Pueden deducirse las accesibilidades a disolvente de aminoácidos de región V a partir de la estructura tridimensional conocida para fragmentos de anticuerpos humanos y murinos. Hay dos etapas generales en el recubrimiento de un sitio de unión a antígeno murino. Inicialmente, se comparan las FR de los dominios variables de una molécula de anticuerpo de interés con secuencias de FR correspondientes de dominios variables humanos a partir de las fuentes identificadas anteriormente. Entonces se comparan las regiones V humanas más homólogas residuo a residuo con aminoácidos murinos correspondientes. Se sustituyen los residuos en la FR murina que se diferencian del homólogo humano por los residuos presentes en el resto humano utilizando técnicas recombinantes bien conocidas en la materia. Sólo se lleva a cabo el cambio de residuos con restos que están al menos parcialmente expuestos (accesibles a disolvente), y se tiene cuidado en la sustitución de residuos de aminoácido que pueden presentar un efecto significativo sobre la estructura terciaria de dominios de región V, tales como prolina, glicina y aminoácidos cargados.

Por tanto, de esta manera, los sitios de unión a antígeno murinos "recubiertos" resultantes están diseñados para conservar los residuos de CDR murinos, los residuos sustancialmente adyacentes a las CDR, los residuos identificados como enterrados o principalmente enterrados (no accesibles a disolvente), los residuos que se cree que participan en contactos no covalentes (por ejemplo, electrostáticos e hidrófobos) entre dominios de cadena pesada y ligera, y los residuos de regiones estructurales conservadas de las FR que se cree que influyen en las estructuras terciarias "canónicas" de los bucles de CDR. Entonces se utilizan estos criterios de diseño para preparar secuencias de nucleótidos recombinantes que combinan las CDR de la cadena tanto pesada como ligera de un sitio de unión a antígeno murino en FR que aparecen en seres humanos que pueden utilizarse para transfectar células de mamífero para la expresión de anticuerpos humanos recombinantes que muestran la especificidad de antígeno de la molécula de anticuerpo murina.

En otra realización de la invención, pueden acoplarse anticuerpos monoclonales de la presente invención a uno o más agentes terapéuticos. Los agentes adecuados con respecto a esto incluyen radionúclidos, inductores de diferenciación, fármacos, toxinas, y derivados de los mismos. Los radionúclidos preferidos incluyen  $^{90}\text{Y}$ ,  $^{123}\text{I}$ ,  $^{125}\text{I}$ ,  $^{131}\text{I}$ ,  $^{186}\text{Re}$ ,  $^{188}\text{Re}$ ,  $^{211}\text{At}$  y  $^{212}\text{Bi}$ . Los fármacos preferidos incluyen metotrexato y análogos de pirimidina y purina. Los inductores de diferenciación preferidos incluyen ésteres de forbol y ácido butírico. Las toxinas preferidas incluyen ricina, abrina, toxina diftérica, toxina del cólera, gelonina, exotoxina de *Pseudomonas*, toxina de *Shigella* y proteína antiviral de fitolaca.

Puede acoplarse un agente terapéutico (por ejemplo, unirse covalentemente) a un anticuerpo monoclonal adecuado o bien directa o bien indirectamente (por ejemplo, a través de un grupo de unión). Una reacción directa entre un agente y un anticuerpo es posible cuando cada uno presenta un sustituyente que puede reaccionar con el otro. Por ejemplo, un grupo nucleófilo, tal como un grupo amino o sulfhidrilo, en uno puede reaccionar con un grupo que contiene carbonilo, tal como un anhídrido o un haluro de ácido, o con un grupo alquilo que contiene un buen grupo saliente (por ejemplo, un haluro) en el otro.

Alternativamente, puede ser deseable acoplar un agente terapéutico y un anticuerpo a través de un grupo de unión. Un grupo de unión puede funcionar como separador para distanciar un anticuerpo de un agente con el fin de interferir con las capacidades de unión. Un grupo de unión también puede servir para aumentar la reactividad química de un sustituyente en un agente o un anticuerpo, y por tanto aumentar la eficacia de acoplamiento. Un aumento en la reactividad química también puede facilitar la utilización de agentes, o grupos funcionales en agentes, que de lo contrario no sería posible.

Resultará evidente para los expertos en la materia que puede emplearse una variedad de reactivos bifuncionales y polifuncionales, tanto homo como heterofuncionales (tales como los descritos en el catálogo de Pierce Chemical Co., Rockford, IL), como grupo de unión. El acoplamiento puede realizarse, por ejemplo, a través de grupos amino, grupos carboxilo, grupos sulfhidrilo o residuos de hidrato de carbono oxidados. Hay numerosas referencias que describen tal metodología, por ejemplo, la patente US n.º 4.671.958, concedida a Rodwell *et al.*

Cuando un agente terapéutico es más potente cuando está libre de la parte de anticuerpo de los inmunoconjugados de la presente invención, puede ser deseable utilizar un grupo de unión que pueda escindirse durante o tras la internalización en una célula. Se han descrito varios grupos de unión escindibles diferentes. Los mecanismos para la liberación intracelular de un agente a partir de estos grupos de unión incluyen escisión mediante reducción de un puente disulfuro (por ejemplo, patente US n.º 4.489.710, concedida a Spitzer), mediante irradiación de un enlace fotolábil (por ejemplo, patente US n.º 4.625.014, concedida a Senter *et al.*), mediante hidrólisis de cadenas laterales de aminoácido derivatizado (por ejemplo, patente US n.º 4.638.045, concedida a Kohn *et al.*), mediante hidrólisis mediada por complemento sérico (por ejemplo, patente US n.º 4.671.958, concedida a Rodwell *et al.*) e hidrólisis catalizada por ácido (por ejemplo, patente US n.º 4.569.789, concedida a Blattler *et al.*).

Puede ser deseable acoplar más de un agente a un anticuerpo. En una realización, se acoplan múltiples moléculas de un agente a una molécula de anticuerpo. En otra realización, puede acoplarse más de un tipo de agente a un anticuerpo. Independientemente de la realización particular, pueden prepararse inmunoconjugados con más de un agente de una variedad de maneras. Por ejemplo, puede acoplarse más de un agente directamente a una molécula

de anticuerpo, o pueden utilizarse ligadores que proporcionan múltiples sitios para la fijación. Alternativamente, puede utilizarse un portador.

Un portador puede llevar los agentes de una variedad de maneras, incluyendo formación de enlaces covalentes o bien directamente o bien a través de un grupo de unión. Los portadores adecuados incluyen proteínas tales como albúminas (por ejemplo, patente US n.º 4.507.234, concedida a Kato *et al.*), péptidos y polisacáridos tales como aminodextrano (por ejemplo, patente US n.º 4.699.784, concedida a Shih *et al.*). Un portador también puede llevar un agente mediante formación de enlaces no covalentes o mediante encapsulación, tal como dentro de una vesícula de liposoma (por ejemplo, patentes US n.ºs 4.429.008 y 4.873.088). Los portadores específicos para agentes de radionúclido incluyen moléculas pequeñas radiohalogenadas y compuestos quelantes. Por ejemplo, la patente US n.º 4.735.792 da a conocer moléculas pequeñas radiohalogenadas representativas y su síntesis. Puede formarse un quelato de radionúclido a partir de compuestos quelantes incluyendo aquellos que contienen átomos de nitrógeno y de azufre como átomos donadores para la unión al radionúclido de metal, u óxido de metal. Por ejemplo, la patente US n.º 4.673.562, concedida a Davison *et al.* da a conocer compuestos quelantes representativos y su síntesis.

#### Composiciones de células T

La presente invención, en otro aspecto, proporciona células T específicas para un polipéptido bacteriano dado a conocer en la presente memoria, o para una variante o derivado del mismo. Tales células pueden prepararse *ex vivo*, utilizando procedimientos convencionales. Por ejemplo, pueden aislarse células T de biopsias, médula ósea, sangre periférica o una fracción de médula ósea o sangre periférica de un paciente, utilizando un sistema de separación de células disponible comercialmente, tal como el sistema Isolex™, disponible de Nexell Therapeutics, Inc. (Irvine, CA; véanse también la patente US n.º 5.240.856; patente US n.º 5.215.926; los documentos WO 89/06280; WO 91/16116 y WO 92/07243). En una realización particular de la presente invención, pueden aislarse células T a partir de muestras de linfocitos intraepiteliales (LIE) o linfocitos de la lámina propia (LLP) que se originan de biopsias de colon. Los individuos expertos en la materia reconocerán fácilmente que hay numerosas metodologías para aislar LIE y LLP (por ejemplo, procedimientos descritos en Christ, A. D., S. P. Colgan, S. P. Balk, R. S. Blumberg. 1997. *Immunol. Lett.* 58:159; Boll G, Reimann J. *Scand J Immunol* agosto de 1995 ;42(2):191-201). En determinados aspectos, pueden derivarse células T de cultivos o líneas de células de mamíferos no humanos, humanos relacionadas o no relacionadas.

Pueden estimularse células T con un polipéptido, polinucleótido que codifica para un polipéptido y/o una célula presentadora de antígeno (CPA) que expresa tal polipéptido. Tal estimulación se realiza en condiciones y durante un tiempo suficientes para permitir la generación de células T que son específicas para el polipéptido de interés. Preferentemente, un polipéptido bacteriano o polinucleótido de la invención está presente dentro de un vehículo de administración, tal como una microesfera, para facilitar la generación de células T específicas.

Se considera que las células T son específicas para un polipéptido de la presente invención si las células T proliferan específicamente, secretan citocinas o destruyen células diana recubiertas con el polipéptido o que expresan un gen que codifica para el polipéptido. Puede evaluarse la especificidad de células T utilizando cualquiera de una variedad de técnicas convencionales. Por ejemplo, dentro de un ensayo de liberación de cromo o ensayo de proliferación, un índice de estimulación con un aumento de más de dos veces de la lisis y/o la proliferación, en comparación con controles negativos, indica especificidad de células T. Tales ensayos pueden realizarse, por ejemplo, tal como se describe en Chen *et al.*, *Cancer Res.* 54:1065-1070, 1994. Alternativamente, la detección de la proliferación de células T puede lograrse mediante una variedad de técnicas conocidas. Por ejemplo, puede detectarse la proliferación de células T midiendo una tasa aumentada de síntesis de ADN (por ejemplo, marcando por impulsos cultivos de células T con timidina tritiada y midiendo la cantidad de timidina tritiada incorporada en el ADN). El contacto con un polipéptido bacteriano (100 ng/ml - 100 µg/ml, preferentemente 200 ng/ml - 25 µg/ml) durante 3 - 7 días dará normalmente como resultado un aumento de al menos dos veces de la proliferación de las células T. El contacto tal como se describió anteriormente durante 2-3 horas debe dar como resultado la activación de las células T, según se mide utilizando ensayos de citocinas convencionales en los que un aumento de dos veces del nivel de liberación de citocina (por ejemplo, TNF o IFN-γ) es indicativo de activación de células T (véase Coligan *et al.*, *Current Protocols in Immunology*, vol. 1, Wiley Interscience (Greene 1998)). Las células T que se han activado en respuesta a un polipéptido bacteriano, polinucleótido o CPA que expresa polipéptido pueden ser CD4<sup>+</sup> y/o CD8<sup>+</sup>. Pueden expandirse células T específicas para polipéptidos bacterianos utilizando técnicas convencionales. Dentro de realizaciones preferidas, las células T se derivan de un paciente, un donante relacionado o un donante no relacionado, y se administran al paciente tras la estimulación y expansión.

Para fines terapéuticos, pueden expandirse en número células T CD4<sup>+</sup> o CD8<sup>+</sup> que proliferan en respuesta a un polipéptido bacteriano, polinucleótido o CPA *in vitro* o *in vivo*. La proliferación de tales células T *in vitro* puede lograrse de una variedad de maneras. Por ejemplo, las células T pueden volver a exponerse a un polipéptido bacteriano, o un péptido corto correspondiente a una parte inmunogénica de tal polipéptido, con o sin la adición de factores de crecimiento de células T, tales como interleucina 2, y/o células estimulantes que sintetizan un polipéptido bacteriano. Alternativamente, pueden expandirse en número una o más células T que proliferan en presencia del polipéptido bacteriano mediante clonación. En la materia se conocen bien procedimientos para clonar células, e incluyen dilución limitante.

En determinadas realizaciones, pueden ser deseables células T que producen citocinas antiinflamatorias. Tales citocinas pueden incluir, pero no se limitan a, 10 (IL-10), interferón $\gamma$  (IFN- $\gamma$ ), interleucina 4 (IL-4), interleucina 12 (IL-12), factor de crecimiento transformante beta (TGF $\beta$ ) e interleucina 18 (IL-18). En determinadas realizaciones, una respuesta antiinflamatoria está mediada por células cooperadoras T CD4<sup>+</sup>.

Composiciones de receptor de células T

El receptor de células T (TCR) consiste en 2 cadenas de polipéptido altamente variables, diferentes, denominadas cadenas  $\alpha$  y  $\beta$  de receptor de células T, que están unidas mediante un puente disulfuro (Janeway, Travers, Walport. Immunobiology. Cuarta ed., 148-159. Elsevier Science Ltd/Garland Publishing. 1999). El heterodímero  $\alpha/\beta$  forma complejos con las cadenas CD3 no variables en la membrana celular. Este complejo reconoce péptidos antigénicos específicos unidos a moléculas de MHC. La enorme diversidad de especificidades de TCR se genera de manera muy similar a la diversidad de inmunoglobulinas, mediante reordenación génica somática. Los genes de la cadena  $\beta$  contienen más de 50 segmentos variables (V), 2 de diversidad (D), más de 10 de unión (J), y 2 segmentos de región constante (C). Los genes de cadena  $\alpha$  contienen más de 70 segmentos V, y más de 60 segmentos J pero ningún segmento D, así como un segmento C. Durante el desarrollo de células T en el timo, se produce la reordenación de genes de D a J de la cadena  $\beta$ , seguida por la reordenación del segmento génico V al DJ. Este exón VDJ $\beta$  funcional se transcribe y se somete a corte y empalme para unirse a un C $\beta$ . Para la cadena  $\alpha$ , un segmento génico V $\alpha$  se reordena a un segmento génico J $\alpha$  para crear un exón funcional que entonces se transcribe y se somete a corte y empalme con C $\alpha$ . Se aumenta adicionalmente la diversidad durante el proceso de recombinación mediante la adición al azar de nucleótidos P y N entre los segmentos V, D y J de la cadena  $\beta$  y entre los segmentos V y J en la cadena  $\alpha$  (Janeway, Travers, Walport. Immunobiology. Cuarta ed., 98 y 150. Elsevier Science Ltd/Garland Publishing. 1999).

En la presente memoria se dan a conocer TCR específicos para un polipéptido dado a conocer en la presente memoria, o para una variante o derivado del mismo. Se dan a conocer secuencias de aminoácidos y polinucleótido para las regiones de unión V-J o V-D-J o partes de las mismas para las cadenas alfa y beta del receptor de células T que reconoce polipéptidos bacterianos descritos en la presente memoria. En general, esta descripción se refiere a receptores de células T que reconocen o se unen a polipéptidos bacterianos presentados en el contexto de MHC. Preferentemente los antígenos bacterianos reconocidos por los receptores de células T comprenden un polipéptido de la presente invención. Por ejemplo, puede aislarse ADNc que codifica para un TCR específico para un péptido bacteriano a partir de células T específicas para un polipéptido bacteriano utilizando técnicas de ADN recombinante y de biología molecular convencionales.

Se dan a conocer receptores de células T o análogos de los mismos que presentan sustancialmente la misma función o actividad que los receptores de células T que reconocen o se unen a polipéptidos bacterianos. Tales receptores incluyen, pero no se limitan a, un fragmento del receptor, o un mutante de sustitución, adición o delección de un receptor de células T descrito en la presente memoria. Esta descripción también abarca polipéptidos o péptidos que son sustancialmente homólogos a los receptores de células T descritos en la presente memoria o que conservan sustancialmente la misma actividad. El término "análogo" incluye cualquier proteína o polipéptido que presenta una secuencia de residuos de aminoácido sustancialmente idéntica a los receptores de células T proporcionados en la presente memoria en las que uno o más residuos, preferentemente no más de 5 residuos, más preferentemente no más de 25 residuos se han sustituido de manera conservativa por un residuo funcionalmente similar y que presenta los aspectos funcionales del receptor de células T tal como se describe en la presente memoria.

Pueden transfectarse células huésped de mamífero adecuadas, por ejemplo, células T no específicas, con un polinucleótido que codifica para TCR específicos para un polipéptido descrito en la presente memoria, haciendo así que la célula huésped sea específica para el polipéptido. Las cadenas  $\alpha$  y  $\beta$  del TCR pueden contenerse en vectores de expresión separados o alternativamente, en un único vector de expresión que también contiene un sitio de entrada a ribosoma interno (IRES) para la traducción independiente del extremo del gen en sentido 3' del IRES. Dichas células huésped que expresan TCR específicos para el polipéptido pueden utilizarse, por ejemplo, para inmunoterapia adoptiva de EIL tal como se comenta adicionalmente a continuación.

Pueden utilizarse TCR clonados específicos para un polipéptido mencionado en la presente memoria en un kit para el diagnóstico de EIL. Por ejemplo, puede utilizarse la secuencia de ácido nucleico o partes de la misma, de TCR específicos de bacterias como sondas o cebadores para la detección de expresión de los genes reordenados que codifican para el TCR específico en una muestra biológica. Por tanto, también se da a conocer un ensayo para detectar ARN mensajero o ADN que codifica para el TCR específico para un polipéptido.

Composiciones farmacéuticas

En realizaciones adicionales, la presente invención se refiere a una formulación de una o más de las composiciones de polinucleótido, polipéptido, células T y/o anticuerpos dadas a conocer en la presente memoria en portadores farmacéuticamente aceptables para su administración a una célula o un animal, o bien sola o bien en combinación

con una o más de otras modalidades de terapia.

Se entenderá que, si se desea, también puede administrarse una composición tal como se da a conocer en la presente memoria en combinación con otros agentes, tales como, por ejemplo, otras proteínas o polipéptidos o diversos agentes farmacéuticamente activos. De hecho, no hay prácticamente ningún límite en cuanto a otros componentes que también pueden incluirse, siempre que los agentes adicionales no provoquen un efecto adverso significativo tras el contacto con tejidos huésped o células diana. Por tanto, las composiciones pueden administrarse junto con varios otros agentes según se requiera en el caso particular. Tales composiciones pueden purificarse a partir de células huésped u otras fuentes biológicas, o alternativamente pueden sintetizarse químicamente tal como se describe en la presente memoria. Igualmente, tales composiciones pueden comprender además composiciones de ARN o ADN sustituidos o derivatizados.

Por tanto, en otro aspecto de la presente invención, se proporcionan composiciones farmacéuticas que comprenden una o más de las composiciones de polinucleótido, polipéptido, anticuerpos y/o CPA descritas en la presente memoria en combinación con un portador fisiológicamente aceptable. En determinadas realizaciones preferidas, las composiciones farmacéuticas de la invención comprenden composiciones de polinucleótido y/o de polipéptido inmunogénicas de la invención para su utilización en aplicaciones de vacuna profilácticas y terapéuticas. La preparación de vacunas se describe de manera general, por ejemplo, en M.F. Powell y M.J. Newman, eds., "Vaccine Design (the subunit and adjuvant approach)", Plenum Press (NY, 1995). Generalmente, tales composiciones comprenderán una o más composiciones de polinucleótido y/o polipéptido de la presente invención en combinación con uno o más inmunoestimulantes.

Resultará evidente que cualquiera de las composiciones farmacéuticas tal como se describen en la presente memoria puede contener sales farmacéuticamente aceptables de los polinucleótidos y polipéptidos de la invención. Tales sales pueden prepararse, por ejemplo, a partir de bases no tóxicas farmacéuticamente aceptables, incluyendo bases orgánicas (por ejemplo, sales de aminas primarias, secundarias y terciarias y aminoácidos básicos) y bases inorgánicas (por ejemplo, sales de sodio, potasio, litio, amonio, calcio y magnesio).

En otra realización, las composiciones inmunogénicas ilustrativas, por ejemplo, composiciones de vacuna, de la presente invención comprenden ADN que codifica para uno o más de los polipéptidos tal como se describió anteriormente, de tal manera que el polipéptido se genera *in situ*. Tal como se indicó anteriormente, el polinucleótido puede administrarse dentro de cualquiera de una variedad de sistemas de administración conocidos por los expertos habituales en la materia. De hecho, en la materia se conocen bien numerosas técnicas de inserción génica, tales como las descritas por Rolland, Crit. Rev. Therap. Drug Carrier Systems 15:143-198, 1998, y referencias citadas en el mismo. Evidentemente, los sistemas de expresión de polinucleótido apropiados contendrán las secuencias reguladoras de ADN necesarias para la expresión en un paciente (tales como un promotor y señal de terminación adecuados). Alternativamente, los sistemas de administración bacterianos pueden implicar la administración de una bacteria (tal como *Bacillus-Calmette-Guerrin*) que expresa una parte inmunogénica del polipéptido sobre su superficie celular o secreta un epítipo de este tipo.

Por tanto, en determinadas realizaciones, se introducen polinucleótidos que codifican para polipéptidos inmunogénicos descritos en la presente memoria en células huésped de mamífero adecuadas para la expresión utilizando cualquiera de varios sistemas basados en virus conocidos. En una realización ilustrativa, los retrovirus proporcionan una plataforma conveniente y eficaz para sistemas de inserción génica. Puede insertarse una secuencia de nucleótidos seleccionada que codifica para un polipéptido de la presente invención en un vector y empaquetarse en partículas retrovirales utilizando técnicas conocidas en la materia. Entonces puede aislarse el virus recombinante y administrarse a un sujeto. Se han descrito varios sistemas retrovirales ilustrativos (por ejemplo, la patente US n.º 5.219.740; Miller y Rosman (1989) BioTechniques 7:980-990; Miller, A. D. (1990) Human Gene Therapy 1:5-14; Scarpa *et al.* (1991) Virology 180: 849-852; Burns *et al.* (1993) Proc. Natl. Acad. Sci. USA 90:8033-8037; y Boris-Lawrie y Temin (1993) Cur. Opin. Genet. Develop. 3:102-109).

Además, también se han descrito varios sistemas basados en adenovirus ilustrativos. Al contrario que los retrovirus que se integran en el genoma huésped, los adenovirus persisten de manera extracromosómica minimizando así los riesgos asociados con mutagénesis por inserción (Haj-Ahmad y Graham (1986) J. Virol. 57:267-274; Bett *et al.* (1993) J. Virol. 67: 5911-5921; Mittereder *et al.* (1994) Human Gene Therapy 5:717-729; Seth *et al.* (1994) J. Virol. 68:933-940; Barr *et al.* (1994) Gene Therapy 1:51-58; Berkner, K. L. (1988) BioTécnicas 6:616-629; y Rich *et al.* (1993) Human Gene Therapy 4:461-476).

También se han desarrollado diversos sistemas de vectores de virus adenoasociados (AAV) para la administración de polinucleótidos. Pueden construirse fácilmente vectores de AAV utilizando técnicas bien conocidas en la materia. Véanse, por ejemplo, las patentes US n.ºs 5.173.414 y 5.139.941; las publicaciones internacionales n.ºs WO 92/01070 y WO 93/03769; Lebkowski *et al.* (1988) Molec. Cell. Biol. 8: 3988-3996; Vincent *et al.* (1990) Vaccines 90 (Cold Spring Harbor Laboratory Press); Carter, B. J. (1992) Current Opinion in Biotechnology 3:533-539; Muzyczka, N. (1992) Current Topics in Microbiol. and Immunol. 158:97-129; Kotin, R. M. (1994) Human Gene Therapy 5:793-801; Shelling y Smith (1994) Gene Therapy 1:165-169; y Zhou *et al.* (1994) J. Exp. Med. 179:1867-1875.

Los vectores virales adicionales útiles para administrar los polinucleótidos que codifican para polipéptidos de la presente invención mediante transferencia génica incluyen aquellos derivados de la familia de virus de la viruela, tales como virus vaccinia y virus de la viruela aviar. A modo de ejemplo, pueden construirse recombinantes de virus vaccinia que expresan las moléculas novedosas de la siguiente manera. En primer lugar se inserta el ADN que codifica para un polipéptido en un vector apropiado de modo que queda adyacente a un promotor de vaccinia y a secuencias de ADN de vaccinia flanqueantes, tales como la secuencia que codifica para timidina cinasa (TK). Entonces se utiliza este vector para transfectar células que se infectan simultáneamente con vaccinia. La recombinación homóloga sirve para insertar el promotor de vaccinia más el gen que codifica para el polipéptido de interés en el genoma viral. El recombinante TK.sup.(-) resultante puede seleccionarse cultivando las células en presencia de 5-bromodesoxiuridina y recogiendo placas de lisis virales resistentes a la misma.

Puede utilizarse convenientemente un sistema de infección/transfección basado en vaccinia para proporcionar expresión o coexpresión inducible, transitoria, de uno o más polipéptidos descritos en la presente memoria en células huésped de un organismo. En este sistema particular, en primer lugar se infectan células *in vitro* con un recombinante de virus vaccinia que codifica para la ARN polimerasa del bacteriófago T7. Esta polimerasa presenta especificidad exquisita porque sólo transcribe moldes que llevan promotores de T7. Tras la infección, se transfectan células con el polinucleótido o polinucleótidos de interés, impulsado por un promotor de T7. La polimerasa expresada en el citoplasma del recombinante de virus vaccinia transcribe el ADN transfectado en ARN, que entonces se traduce en polipéptido por la maquinaria de traducción del huésped. El procedimiento proporciona producción citoplasmática, transitoria, a alto nivel, de grandes cantidades de ARN y sus productos de traducción. Véanse, por ejemplo, Elroy-Stein y Moss, Proc. Natl. Acad. Sci. USA (1990) 87:6743-6747; Fuerst *et al.* Proc. Natl. Acad. Sci. USA (1986) 83:8122-8126.

Alternativamente, también pueden utilizarse virus de la viruela aviar, tales como los virus de la viruela de la gallina y de la viruela del canario, para administrar las secuencias codificantes de interés. Se sabe que virus de la viruela aviar recombinantes, que expresan inmunógenos de patógenos de mamífero, confieren inmunidad protectora cuando se administran a especies no aviares. La utilización de un vector de la viruela aviar es particularmente deseable en seres humanos y otras especies de mamíferos ya que los miembros del género de la viruela aviar sólo pueden replicarse productivamente en especies aviares sensibles y por tanto no son infecciosos en células de mamífero. En la materia se conocen procedimientos para producir virus de la viruela aviar recombinantes y emplean recombinación genética, tal como se describió anteriormente con respecto a la producción de virus vaccinia. Véanse, por ejemplo, los documentos WO 91/12882; WO 89/03429 y WO 92/03545.

También puede utilizarse cualquiera de varios vectores de alfavirus para la administración de composiciones de polinucleótido de la presente invención, tales como los vectores descritos en las patentes US n.ºs 5.843.723; 6.015.686; 6.008.035 y 6.015.694. También pueden utilizarse determinados vectores basados en la encefalitis equina de Venezuela (EEV), de los cuales pueden encontrarse ejemplos ilustrativos en las patentes US n.ºs 5.505.947 y 5.643.576.

Además, también pueden utilizarse vectores de conjugados moleculares, tales como los vectores quiméricos de adenovirus descritos en Michael *et al.* J. Biol. Chem. (1993) 268:6866-6869 y Wagner *et al.* Proc. Natl. Acad. Sci. USA (1992) 89:6099-6103, para la inserción génica según la invención.

Puede encontrarse información ilustrativa adicional sobre estos y otros sistemas de administración basados en virus conocidos, por ejemplo, en Fisher-Hoch *et al.*, Proc. Natl. Acad. Sci. USA 86:317-321, 1989; Flexner *et al.*, Ann. N.Y. Acad. Sci. 569: 86-103, 1989; Flexner *et al.*, Vacuna 8:17-21, 1990; patentes US n.ºs 4.603.112, 4.769.330 y 5.017.487; documento WO 89/01973; patente US n.º 4.777.127; documentos GB 2.200.651; EP 0.345.242; WO 91/02805; Berkner, Biotechniques 6:616-627, 1988; Rosenfeld *et al.*, Science 252:431-434, 1991; Kolls *et al.*, Proc. Natl. Acad. Sci. USA 91:215-219, 1994; Kass- Eisler *et al.*, Proc. Natl. Acad. Sci. USA 90:11498-11502, 1993; Guzman *et al.*, Circulación 88:2838-2848, 1993; y Guzman *et al.*, Cir. Res. 73:1202-1207, 1993.

En determinadas realizaciones, puede integrarse un polinucleótido en el genoma de una célula diana. Esta integración puede ser en la ubicación y orientación específicas mediante recombinación homóloga (sustitución génica) o puede integrarse en una ubicación al azar, no específica (aumento génico). En realizaciones aún adicionales, el polinucleótido puede mantenerse de manera estable en la célula como un segmento de ADN episomal separado. Tales segmentos de polinucleótido o "episomas" codifican para secuencias suficientes para permitir el mantenimiento y la replicación independientes de, o en sincronización con, el ciclo de la célula huésped. La manera en la que se administra el constructo de expresión a una célula y dónde en la célula permanece el polinucleótido depende del tipo de constructo de expresión empleado.

En otra realización de la invención, un polinucleótido se administra/suministra como ADN "desnudo", por ejemplo tal como se describe en Ulmer *et al.*, Science 259:1745-1749, 1993 y se revisa por Cohen, Science 259:1691-1692, 1993. La captación de ADN desnudo puede aumentarse recubriendo el ADN sobre perlas biodegradables, que se transportan eficazmente al interior de las células.

En todavía otra realización, puede administrarse una composición de la presente invención mediante un enfoque de

- bombardeo de partículas, muchos de los cuales se han descrito. En un ejemplo ilustrativo, puede lograrse la aceleración de partículas impulsada por gas con dispositivos tales como los fabricados por Powderject Pharmaceuticals PLC (Oxford, RU) y Powderject Vaccines Inc. (Madison, WI), algunos ejemplos de los cuales se describen en las patentes US n.ºs 5.846.796; 6.010.478; 5.865.796; 5.584.807; y la patente EP n.º 0500 799. Este enfoque ofrece un enfoque de administración sin agujas en el que se acelera una formulación en polvo seco de partículas microscópicas, tales como partículas de polinucleótido o polipéptido, hasta alta velocidad dentro de un chorro de gas helio generado por un dispositivo portátil, impulsando las partículas al interior de un tejido diana de interés.
- En una realización relacionada, otros dispositivos y procedimientos que pueden ser útiles para la inyección sin agujas impulsada por gas de composiciones de la presente invención incluyen los proporcionados por Bioject, Inc. (Portland, OR), algunos ejemplos de los cuales se describen en las patentes US n.ºs 4.790.824; 5.064.413; 5.312.335; 5.383.851; 5.399.163; 5.520.639 y 5.993.412.
- Según otra realización, las composiciones farmacéuticas tal como se describen en la presente memoria comprenderán uno o más inmunoestimulantes además de las composiciones de polinucleótido, polipéptido, anticuerpos, células T y/o CPA inmunogénicas de esta invención. Un inmunoestimulante se refiere esencialmente a cualquier sustancia que aumenta o potencia una respuesta inmunitaria (mediada por anticuerpos y/o células) frente a un antígeno exógeno. Un tipo preferido de inmunoestimulante comprende un adyuvante. Muchos adyuvantes contienen una sustancia diseñada para proteger al antígeno frente a un catabolismo rápido, tal como hidróxido de aluminio o aceite mineral, y un estimulante de respuestas inmunitarias, tal como lípido A, proteínas derivadas de *Bordetella pertussis* o *Mycobacterium tuberculosis*. Determinados adyuvantes están disponibles comercialmente tales como, por ejemplo, adyuvante incompleto y adyuvante completo de Freund (Difco Laboratories, Detroit, MI); adyuvante 65 de Merck (Merck and Company, Inc., Rahway, NJ); AS-2 (SmithKline Beecham, Filadelfia, PA); sales de aluminio tales como gel de hidróxido de aluminio (alum) o fosfato de aluminio; sales de calcio, hierro o zinc; una suspensión insoluble de tirosina acilada; azúcares acilados; polisacáridos derivados catiónica o aniónicamente; polifosfacenos; microesferas biodegradables; monofosforil lípido A y Quil A. También pueden utilizarse como adyuvantes citocinas, tales como GM-CSF, interleucina 2, 7, 12, y otros factores de crecimiento similares.
- Dentro de determinadas realizaciones de la invención, la composición de adyuvante es preferentemente una que induce una respuesta inmunitaria antiinflamatoria (mediada por anticuerpos o células). Por consiguiente, se preferirán altos niveles de citocinas antiinflamatorias (las citocinas antiinflamatorias pueden incluir, pero no se limitan a, interleucina 4 (IL-4), interleucina 5 (IL-5), interleucina 10 (IL-10) y factor de crecimiento transformante beta (TGF(β)). En determinadas realizaciones, una respuesta antiinflamatoria estará mediada por células cooperadoras T CD4<sup>+</sup>. Se ha sugerido que la flagelina bacteriana actúa como adyuvante (McSorley *et al.*, J. Immunol. 169:3914-19, 2002). Dentro de una realización de la invención, las proteínas de flagelina dadas a conocer en la presente memoria pueden utilizarse en composiciones de adyuvante.
- Dentro de otras realizaciones, los adyuvantes utilizados junto con las composiciones de la presente invención aumentan la sensibilidad a lipopolisacárido (LPS). Los adyuvantes ilustrativos incluyen, pero no se limitan a, monofosforil lípido A (MPL), 4-fosfatos de aminoalquil-glucosaminida (AGP), incluyendo RC-512, RC-522, RC-527, RC-529, RC-544 y RC-560 (Corixa, Hamilton, MT) y otros AGP tales como los descritos en las solicitudes de patente US en tramitación junto con la presente con n.ºs de serie 08/853.826 y 09/074.720.
- Dentro de otras realizaciones de la invención, la composición de adyuvante es una que induce una respuesta inmunitaria predominantemente de tipo Th1. Altos niveles de citocinas de tipo Th1 (por ejemplo, IFN-γ, TNFα, IL-2 y IL-12) tienden a favorecer la inducción de respuestas inmunitarias mediadas por células frente a un antígeno administrado. En cambio, altos niveles de citocinas de tipo Th2 (por ejemplo, IL-4, IL-5, IL-6 y IL-10) tienden a favorecer la inducción de respuestas inmunitarias humorales. Tras la aplicación de una vacuna tal como se proporciona en la presente memoria, un paciente soportará una respuesta inmunitaria que incluye respuestas de tipo Th1 y Th2. Dentro de una realización preferida, en la que una respuesta es predominantemente de tipo Th1, el nivel de citocinas de tipo Th1 aumentará en mayor medida que el nivel de citocinas de tipo Th2. Los niveles de estas citocinas pueden evaluarse fácilmente utilizando ensayos convencionales. Para una revisión de las familias de citocinas, véase Mosmann y Coffman, Ann. Rev. Immunol. 7: 145-173, 1989. Alternativamente, en una realización relacionada, en la que una respuesta preferida es predominantemente de tipo Th2, el nivel de citocinas de tipo Th2 aumentará en mayor medida que el nivel de citocinas de tipo Th1. De nuevo, los niveles de estas citocinas pueden evaluarse fácilmente utilizando ensayos convencionales.
- Determinados adyuvantes preferidos para provocar una respuesta predominantemente de tipo Th1 incluyen, por ejemplo, una combinación de monofosforil lípido A, preferentemente monofosforil lípido A 3-des-O-acilado, junto con una sal de aluminio. Adyuvantes MPL® están disponibles de Corixa Corporation (Seattle, WA; véanse, por ejemplo, las patentes US n.ºs 4.436.727; 4.877.611; 4.866.034 y 4.912.094). Oligonucleótidos que contienen CpG (en los que el dinucleótido CpG no está metilado) también inducen una respuesta predominantemente Th1. Tales oligonucleótidos se conocen bien y se describen, por ejemplo, en los documentos WO 96/02555, WO 99/33488 y las patentes US n.ºs 6.008.200 y 5.856.462. También se describen secuencias de ADN inmunoestimulantes, por ejemplo, por Sato *et al.*, Science 273:352, 1996. Otro adyuvante preferido comprende una saponina, tal como Quil A,

o derivados de la misma, incluyendo QS21 y QS7 (Aquila Biopharmaceuticals Inc., Framingham, MA); escina; digitonina; o saponinas de *Gypsophila* o *Chenopodium quinoa*. Otras formulaciones preferidas incluyen más de una saponina en las combinaciones de adyuvante de la presente invención, por ejemplo combinaciones de al menos dos del siguiente grupo que comprende QS21, QS7, Quil A,  $\beta$ -escina o digitonina.

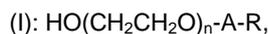
Alternativamente las formulaciones de saponina pueden combinarse con vehículos de vacuna compuestos por quitosano u otros polímeros policatiónicos, partículas de polilactida y polilactida-co-glicolida, matriz polimérica a base de poli-N-acetil-glucosamina, partículas compuestas por polisacáridos o polisacáridos químicamente modificados, liposomas y partículas basadas en lípidos, partículas compuestas por monoésteres de glicerol, etc. Las saponinas también pueden formularse en presencia de colesterol para formar estructuras particuladas tales como liposomas o ISCOM. Además, las saponinas pueden formularse junto con un éster o éter de polioxietileno, o bien en una disolución o suspensión no particulada, o bien en una estructura particulada tal como un liposoma paucilamelar o ISCOM. Las saponinas también pueden formularse con excipientes tales como Carbopol<sup>R</sup> para aumentar la viscosidad, o pueden formularse en una forma de polvo seco con un excipiente en polvo tal como lactosa.

En una realización preferida, el sistema de adyuvante incluye la combinación de un monofosforil lípido A y un derivado de saponina, tal como la combinación de QS21 y adyuvante 3D-MPL®, tal como se describe en el documento WO 94/00153, o una composición menos reactogénica en la que QS21 se extingue con colesterol, tal como se describe en el documento WO 96/33739. Otras formulaciones preferidas comprenden una emulsión de aceite en agua y tocoferol. Otra formulación de adyuvante particularmente preferida que emplea QS21, adyuvante 3D-MPL® y tocoferol en una emulsión de aceite en agua se describe en el documento WO 95/17210.

Otro sistema de adyuvante potenciado implica la combinación de un oligonucleótido que contiene CpG y un derivado de saponina, particularmente la combinación de CpG y QS21 se da a conocer en el documento WO 00/09159. Preferentemente la formulación comprende adicionalmente una emulsión de aceite en agua y tocoferol.

Los adyuvantes ilustrativos adicionales para su utilización en las composiciones farmacéuticas de la invención incluyen Montanide ISA 720 (Seppic, Francia), SAF (Chiron, California, Estados Unidos), ISCOMS (CSL), MF-59 (Chiron), la serie de adyuvantes SBAS (por ejemplo, SBAS-2 o SBAS-4, disponible de SmithKline Beecham, Rixensart, Bélgica), Detox (Enhanzyn®) (Corixa, Hamilton, MT), RC-529 (Corixa, Hamilton, MT) y otros 4-fosfatos de aminoalquil-glucosaminida (AGP), tales como los descritos en las solicitudes de patente US en tramitación junto con la presente con n.ºs de serie 08/853.826 y 09/074.720, y adyuvantes de éter de polioxietileno tales como los descritos en el documento WO 99/52549A1.

Otros adyuvantes preferidos incluyen moléculas de adyuvantes de la fórmula general



en la que, n es 1-50, A es un enlace o -C(O)-, R es alquilo C<sub>1-50</sub> o fenil-alquilo C<sub>1-50</sub>.

Una realización de la presente invención consiste en una formulación de vacuna que comprende un éter de polioxietileno de fórmula general (I), en la que n es entre de 1 y 50, preferentemente 4-24, lo más preferentemente 9; el componente R es alquilo C<sub>1-50</sub>, preferentemente C<sub>4</sub>-C<sub>20</sub> y lo más preferentemente alquilo C<sub>12</sub>, y A es un enlace. La concentración de los éteres de polioxietileno debe estar en el intervalo del 0,1-20%, preferentemente desde el 0,1-10%, y lo más preferentemente en el intervalo del 0,1-1%. Se seleccionan éteres de polioxietileno preferidos del siguiente grupo: 9-lauril éter de polioxietileno, 9-estearil éter de polioxietileno, 8-estearil éter de polioxietileno, 4-lauril éter de polioxietileno, 35-lauril éter de polioxietileno y 23-lauril éter de polioxietileno. Se describen éteres de polioxietileno tales como lauril éter de polioxietileno en el índice de Merck (12ª edición: entrada 7717). Estas moléculas de adyuvantes se describen en el documento WO 99/52549.

El éter de polioxietileno según la fórmula general (I) anterior puede combinarse, si se desea, con otro adyuvante. Por ejemplo, una combinación de adyuvante preferida es preferentemente con CpG tal como se describe en la solicitud de patente UK en tramitación GB 9820956.2.

Según otra realización de esta invención, se administra una composición inmunogénica descrita en la presente memoria a un huésped mediante células presentadoras de antígeno (CPA), tales como células dendríticas, macrófagos, células B, monocitos y otras células que pueden diseñarse mediante ingeniería para ser CPA eficaces. Tales células pueden modificarse genéticamente, pero no es necesario, para aumentar la capacidad para presentar el antígeno, para mejorar la activación y/o el mantenimiento de la respuesta de células T, para presentar efectos antibacterianos en sí mismas y/o para ser inmunológicamente compatibles con el receptor (es decir, haplotipo de HLA coincidente). Generalmente pueden aislarse CPA a partir de cualquiera de una variedad de órganos y fluidos biológicos, incluyendo tejidos bacterianos y peribacterianos, y pueden ser células autólogas, alogénicas, singénicas o xenogénicas.

Determinadas realizaciones preferidas de la presente invención utilizan células dendríticas o progenitores de las mismas como células presentadoras de antígeno. Las células dendríticas son CPA altamente potentes (Banchereau

y Steinman, Nature 392:245-251, 1998) y se ha mostrado que son eficaces como adyuvante fisiológico para provocar inmunidad antibacteriana profiláctica o terapéutica (véase Timmerman y Levy, Ann. Rev. Med. 50:507-529, 1999). En general, pueden identificarse células dendríticas basándose en su forma típica (estrellada *in situ*, con prolongaciones citoplasmáticas marcadas (dendríticas) visibles *in vitro*), su capacidad para captar, procesar y presentar antígenos con alta eficacia y su capacidad para activar respuestas de células T indiferenciadas. Evidentemente, pueden diseñarse mediante ingeniería células dendríticas para expresar ligandos o receptores de superficie celular específicos que no se encuentran comúnmente en células dendríticas *in vivo* o *ex vivo*, y tales células dendríticas modificadas se contemplan por la presente invención. Como alternativa a las células dendríticas, pueden utilizarse células dendríticas cargadas con antígeno de vesículas secretadas (denominadas exosomas) dentro de una vacuna (véase Zitvogel *et al.*, Nature Med. 4:594-600, 1998).

Pueden obtenerse células dendríticas y progenitores a partir de sangre periférica, médula ósea, células infiltrantes de tumores, células infiltrantes de tejidos peritumorales, ganglios linfáticos, bazo, piel, sangre del cordón umbilical o cualquier otro tejido o fluido adecuado. Por ejemplo, pueden diferenciarse células dendríticas *ex vivo* añadiendo una combinación de citocinas tales como GM-CSF, IL-4, IL-13 y/o TNF $\alpha$  a cultivos de monocitos recogidos de sangre periférica. Alternativamente, pueden diferenciarse células positivas para CD34 recogidas de sangre periférica, sangre del cordón umbilical o médula ósea en células dendríticas añadiendo al medio de cultivo combinaciones de GM-CSF, IL-3, TNF $\alpha$ , ligando CD40, LPS, ligando flt3 y/u otro(s) compuesto(s) que induce(n) diferenciación, maduración y proliferación de células dendríticas.

Las células dendríticas se clasifican convenientemente como células "inmaduras" y "maduras", lo que permite una manera sencilla de distinguir entre dos fenotipos bien caracterizados. Sin embargo, no debe interpretarse que esta nomenclatura excluye todos los posibles estadios intermedios de diferenciación. Las células dendríticas inmaduras se caracterizan como CPA con una alta capacidad para la captación y el procesamiento de antígenos, lo que se correlaciona con la alta expresión de receptor de Fc $\gamma$  y receptor de manosa. El fenotipo maduro se caracteriza normalmente por una expresión inferior de estos marcadores, pero una alta expresión de moléculas de superficie celular responsables de la activación de células T tales como MHC de clase I y clase II, moléculas de adhesión (por ejemplo, CD54 y CD11) y moléculas coestimulantes (por ejemplo, CD40, CD80, CD86 y 4-1 BB).

Generalmente pueden transfectarse CPA con un polinucleótido de la invención (o parte u otra variante del mismo) de tal manera que el polipéptido codificado, o una parte inmunogénica del mismo, se expresa sobre la superficie celular. Tal transfección puede tener lugar *ex vivo*, y entonces puede utilizarse una composición farmacéutica que comprende tales células transfectadas para fines terapéuticos, tal como se describe en la presente memoria. Alternativamente, puede administrarse un vehículo de inserción génica que selecciona como diana una célula dendrítica u otra célula presentadora de antígeno a un paciente, dando como resultado la transfección que se produce *in vivo*. La transfección *in vivo* y *ex vivo* de células dendríticas, por ejemplo, puede realizarse generalmente utilizando cualquier procedimiento conocido en la materia, tal como los descritos en el documento WO 97/24447, o el enfoque de pistola génica descrito por Mahvi *et al.*, Immunology and cell Biology 75:456-460, 1997. Puede lograrse la carga con antígeno de células dendríticas incubando células dendríticas o células progenitoras con el polipéptido bacteriano, ADN (desnudo o dentro de un vector de plásmido) o ARN; o con virus o bacterias recombinantes que expresan antígeno (por ejemplo, vectores de vaccinia, viruela aviar, adenovirus o lentivirus). Antes de la carga, puede conjugarse covalentemente el polipéptido a una pareja inmunológica que proporciona cooperación de células T (por ejemplo, una molécula de portador). Alternativamente, puede pulsarse una célula dendrítica con una pareja inmunológica no conjugada, por separado o en presencia del polipéptido.

Aunque puede emplearse cualquier portador adecuado conocido por los expertos habituales en la materia en las composiciones farmacéuticas de esta invención, el tipo de portador variará normalmente dependiendo del modo de administración. Pueden formularse composiciones de la presente invención para cualquier manera apropiada de administración, incluyendo por ejemplo, administración tópica, oral, nasal, mucosa, intravenosa, intracraneal, intraperitoneal, subcutánea e intramuscular.

Los portadores para su utilización dentro de tales composiciones farmacéuticas son biocompatibles, y también pueden ser biodegradables. En determinadas realizaciones, la formulación proporciona preferentemente un nivel relativamente constante de liberación de principio activo. Sin embargo, en otras realizaciones puede desearse una tasa más rápida de liberación inmediatamente tras la administración. La formulación de tales composiciones está dentro del nivel de experiencia habitual en la materia utilizando técnicas conocidas. Los portadores ilustrativos útiles con respecto a esto incluyen micropartículas de poli(láctida-co-glicolida), poliacrilato, látex, almidón, celulosa, dextrano y similares. Otros portadores de liberación retrasada ilustrativos incluyen biovectores supramoleculares, que comprenden un núcleo hidrófilo no líquido (por ejemplo, un oligosacárido o polisacárido reticulado) y, opcionalmente, una capa externa que comprende un compuesto anfífilo, tal como un fosfolípido (véase por ejemplo, la patente US n.º 5.151.254 y las solicitudes PCT WO 94/20078, WO/94/23701 y WO 96/06638). La cantidad de principio activo contenida dentro de una formulación de liberación sostenida depende del sitio de implantación, la tasa y la duración esperada de liberación y la naturaleza del estado que va a tratarse o prevenirse.

En otra realización ilustrativa, se emplean microesferas biodegradables (por ejemplo, polilactato-poliglicolato) como portadores para las composiciones de esta invención. Se dan a conocer microesferas biodegradables adecuadas,

por ejemplo, en las patentes US n.<sup>os</sup> 4.897.268; 5.075.109; 5.928.647; 5.811.128; 5.820.883; 5.853.763; 5.814.344. 5.407.609 y 5.942.252. Sistemas de portador de proteína de núcleo de hepatitis B modificada tales como los descritos en el documento WO/99 40934, y referencias citadas en el mismo, también serán útiles para muchas aplicaciones. Otro sistema de portador/administración ilustrativo emplea un portador que comprende complejos  
5 partícula-proteína, tales como los descritos en la patente US n.º 5.928.647, que pueden inducir respuesta de linfocitos T citotóxicos limitadas a clase I en un huésped.

En otra realización ilustrativa, se emplean partículas de núcleo de fosfato de calcio como portadores, adyuvantes de  
10 vacuna o como matrices de liberación controlada para las composiciones de esta invención. Se dan a conocer partículas de fosfato de calcio a modo de ejemplo, por ejemplo, en la solicitud de patente publicada n.º WO/0046147.

Las composiciones farmacéuticas de la invención comprenderán además con frecuencia uno o más tampones (por  
15 ejemplo, solución salina tamponada neutra o solución salina tamponada con fosfato), hidratos de carbono (por ejemplo, glucosa, manosa, sacarosa o dextranos), manitol, proteínas, polipéptidos o aminoácidos tales como glicina, antioxidantes, bacteriostáticos, agentes quelantes tales como EDTA o glutatión, adyuvantes (por ejemplo, hidróxido de aluminio), solutos que hacen que la formulación sea isotónica, hipotónica y débilmente hipertónica con la sangre de un receptor, agentes de suspensión, agentes espesantes y/o conservantes. Alternativamente, pueden formularse composiciones de la presente invención como liofilizado.

Las composiciones farmacéuticas tal como se describen en la presente memoria pueden presentarse en recipientes  
20 de una única dosis o de múltiples dosis, tales como viales o ampollas selladas. Tales recipientes se sellan normalmente de tal manera que se conserva la esterilidad y estabilidad de la formulación hasta su utilización. En general, las formulaciones pueden almacenarse como suspensiones, disoluciones o emulsiones en vehículos aceitosos o acuosos. Alternativamente, puede almacenarse una composición farmacéutica en un estado liofilizado que sólo requiere la adición de un portador líquido estéril inmediatamente antes de su utilización.

El desarrollo de regímenes de dosificación y tratamiento adecuados para utilizar las composiciones particulares  
30 descritas en la presente memoria en una variedad de regímenes de tratamiento, incluyendo por ejemplo, formulación y administración oral, parenteral, intravenosa, intranasal e intramuscular, se conoce bien en la materia, algunos de los cuales se comentan resumidamente a continuación para fines generales de ilustración.

En determinadas aplicaciones, las composiciones farmacéuticas dadas a conocer en la presente memoria pueden  
35 administrarse mediante administración oral a un animal. Como tales, estas composiciones pueden formularse con un diluyente inerte o con un portador comestible asimilable, o pueden encerrarse en cápsulas de gelatina de cubierta dura o blanda, o pueden comprimirse en comprimidos, o pueden incorporarse directamente en la alimentación de la dieta.

Los principios activos pueden incluso incorporarse con excipientes y utilizarse en forma de comprimidos que pueden  
40 ingerirse, comprimidos bucales, trociscos, cápsulas, elixires, suspensiones, jarabes, obleas, y similares (véanse, por ejemplo, Mathiowitz *et al.*, Nature 27 de mayo de 1997; 386(6623):410-4; Hwang *et al.*, Crit Rev Ther Drug Carrier Syst 1998; 15(3):243-84; patente US 5.641.515; patente US 5.580.579 y patente US 5.792.451). Los comprimidos, trociscos, pastillas, cápsulas y similares también pueden contener cualquiera de una variedad de componentes adicionales, por ejemplo, puede añadirse un aglutinante, tal como goma tragacanto, goma arábiga, almidón de maíz o gelatina; excipientes, tales como fosfato de dicalcio; un agente disgregante, tal como almidón de maíz, almidón de patata, ácido algínico y similares; un lubricante, tal como estearato de magnesio; y un agente edulcorante, tal como sacarosa, lactosa o sacarina o un agente aromatizante, tal como aromatizante menta, aceite de menta fresca o cereza. Cuando la forma farmacéutica unitaria es una cápsula, puede contener, además de materiales del tipo anterior, un portador líquido. Varios otros materiales pueden estar presentes como recubrimientos o para modificar de otro modo la forma física de la unidad de dosificación. Por ejemplo, pueden recubrirse comprimidos, pastillas o  
50 cápsulas con laca, azúcar o ambos. Evidentemente, cualquier material utilizado en la preparación de cualquier forma farmacéutica unitaria debe ser farmacéuticamente puro y sustancialmente no tóxico en las cantidades empleadas. Además, los principios activos pueden incorporarse en formulaciones y preparaciones de liberación sostenida.

Normalmente, estas formulaciones contendrán al menos aproximadamente el 0,1% del principio activo o más,  
55 aunque el porcentaje del/de los principio(s) activo(s) puede variarse, evidentemente, y puede ser convenientemente de entre aproximadamente el 1 o el 2% y aproximadamente el 60% o el 70% o más del peso o el volumen de la formulación total. Naturalmente, la cantidad de principio(s) activo(s) que puede prepararse en cada composición terapéuticamente útil es de tal manera que se obtendrá una dosificación adecuada en cualquier dosis unitaria dada del compuesto. Un experto en la materia de preparación de tales formulaciones farmacéuticas contemplará factores  
60 tales como solubilidad, biodisponibilidad, semivida biológica, vía de administración, término de caducidad del producto, así como otras consideraciones farmacológicas, y como tal, puede ser deseable una variedad de dosificaciones y regímenes de tratamiento.

Para la administración oral las composiciones de la presente invención pueden incorporarse alternativamente con  
65 uno o más excipientes en forma de enjuague bucal, dentífrico, comprimido bucal, pulverización oral o formulación administrada por vía oral sublingual. Alternativamente, el principio activo puede incorporarse en una disolución oral

tal como una que contiene borato de sodio, glicerina y bicarbonato de potasio, o dispersarse en un dentífrico, o añadirse en una cantidad terapéuticamente eficaz a una composición que puede incluir agua, aglutinantes, agentes abrasivos, agentes aromatizantes, agentes espumantes y humectantes. Alternativamente las composiciones pueden prepararse en forma de comprimido o disolución que puede colocarse bajo la lengua o disolverse de otro modo en la boca.

En determinadas circunstancias será deseable administrar las composiciones farmacéuticas dadas a conocer en la presente memoria por vía parenteral, intravenosa, intramuscular o incluso intraperitoneal. Tales enfoques los conoce bien el experto en la materia, algunos de los cuales se describen adicionalmente, por ejemplo, en la patente US 5.543.158; la patente US 5.641.515 y la patente US 5.399.363. En determinadas realizaciones, pueden prepararse disoluciones de los principios activos como base libre o sales farmacológicamente aceptables en agua adecuadamente mezclada con un tensioactivo, tal como hidroxipropilcelulosa. También pueden prepararse dispersiones en glicerol, polietilenglicoles líquidos, y mezclas de los mismos en aceites. En condiciones habituales de almacenamiento y utilización, estas preparaciones contendrán generalmente un conservante para impedir el crecimiento de microorganismos.

Las formas farmacéuticas ilustrativas para su utilización inyectable incluyen disoluciones o dispersiones acuosas estériles y polvos estériles para la preparación extemporánea de disoluciones o dispersiones inyectables estériles (por ejemplo, véase la patente US 5.466.468). En todos los casos la forma debe ser estéril y debe ser fluida hasta el grado en que exista fácil jeringabilidad. Debe ser estable en las condiciones de fabricación y almacenamiento y debe protegerse frente a la acción contaminante de microorganismos, tales como bacterias y hongos. El portador puede ser un medio disolvente o de dispersión que contiene, por ejemplo, agua, etanol, poliol (por ejemplo, glicerol, propilenglicol y polietilenglicol líquido, y similares), mezclas adecuadas de los mismos y/o aceites vegetales. Puede mantenerse una fluidez apropiada, por ejemplo, mediante la utilización de un recubrimiento, tal como lecitina, mediante el mantenimiento del tamaño de partícula requerido en el caso de la dispersión y/o mediante la utilización de tensioactivos. La prevención de la acción de microorganismos puede facilitarse mediante diversos agentes antibacterianos y antifúngicos, por ejemplo, parabenos, clorobutanol, fenol, ácido sórbico, timerosal, y similares. En muchos casos, será preferible incluir agentes isotónicos, por ejemplo, azúcares o cloruro de sodio. Puede provocarse la absorción prolongada de las composiciones inyectables mediante la utilización en las composiciones de agentes que retrasan la absorción, por ejemplo, monoestearato de aluminio y gelatina.

En una realización, para la administración parenteral en una disolución acuosa, debe tamponarse adecuadamente la disolución si es necesario y en primer lugar hacer que el diluyente líquido sea isotónico con glucosa o solución salina suficiente. Estas disoluciones acuosas particulares son especialmente adecuadas para la administración intravenosa, intramuscular, subcutánea e intraperitoneal. Con respecto a esto, los expertos en la materia conocerán un medio acuoso estéril que pueda emplearse a la luz de la presente descripción. Por ejemplo, puede disolverse una dosificación en 1 ml de disolución isotónica de NaCl y o bien añadirse a 1000 ml de fluido de hipodermoclasia o bien inyectarse en el sitio de infusión propuesto (véase por ejemplo, "Remington's Pharmaceutical Sciences" 15ª edición, páginas 1035-1038 y 1570-1580). Se producirá necesariamente alguna variación en la dosificación dependiendo del estado del sujeto que está tratándose. Además, para la administración a seres humanos, evidentemente las preparaciones cumplirán preferentemente con las normas de esterilidad, pirogenicidad y de pureza y seguridad generales requeridas por la Oficina de normas biológicas de la FDA.

En otra realización de la invención, las composiciones dadas a conocer en la presente memoria pueden formularse en una forma neutra o de sal. Las sales farmacéuticamente aceptables ilustrativas incluyen sales de adición de ácido (formadas con los grupos amino libres de la proteína) y que se forman con ácidos inorgánicos tales como, por ejemplo, ácidos clorhídrico o fosfórico, o ácidos orgánicos tales como acético, oxálico, tartárico, mandélico, y similares. También pueden derivarse sales formadas con los grupos carboxilo libres a partir de bases inorgánicas tales como, por ejemplo, hidróxidos de sodio, potasio, amonio, calcio o férrico, y bases orgánicas tales como isopropilamina, trimetilamina, histidina, procaína y similares. Tras la formulación, se administrarán disoluciones de una manera compatible con la formulación de dosificación y en una cantidad tal como sea terapéuticamente eficaz.

Los portadores pueden comprender además todos y cada uno de los disolventes, medios de dispersión, vehículos, recubrimientos, diluyentes, agentes antibacterianos y antifúngicos, agentes isotónicos y de retraso de la absorción, tampones, disoluciones de portador, suspensiones, coloides, y similares. La utilización de tales medios y agentes para sustancias farmacéuticamente activas se conoce bien en la materia. Excepto cuando cualquier agente o medio convencional sea incompatible con el principio activo, se contempla su utilización en las composiciones terapéuticas. También pueden incorporarse principios activos complementarios en las composiciones. La frase "farmacéuticamente aceptable" se refiere a entidades moleculares y composiciones que no producen una reacción alérgica o desfavorable similar cuando se administran a un ser humano.

En determinadas realizaciones, las composiciones farmacéuticas pueden administrarse mediante pulverizaciones intranasales, inhalación y/u otros vehículos de administración de aerosol. Se han descrito procedimientos para administrar genes, ácidos nucleicos y composiciones de péptidos directamente a los pulmones mediante pulverizaciones de aerosol nasal, por ejemplo, en la patente US 5.756.353 y la patente US 5.804.212. Asimismo, también se conoce la administración de fármacos utilizando resinas de micropartículas intranasales (Takenaga *et al.*,

J Controlled Release 2 de marzo de 1998; 52(1-2):81-7) y compuestos de lisofosfatidil-glicerol (patente US 5.725.871) en las materias farmacéuticas. Asimismo, la administración de fármacos transmucosa ilustrativa en forma de una matriz de soporte de politetrafluoroetileno se describe en la patente US 5.780.045.

5 En determinadas realizaciones, se utilizan liposomas, nanocápsulas, micropartículas, partículas lipídicas, vesículas, y similares, para la introducción de las composiciones de la presente invención en células/organismos huésped adecuados. En particular, las composiciones de la presente invención pueden formularse para su administración encapsuladas en una partícula lipídica, un liposoma, una vesícula, una nanoesfera o una nanopartícula o similares. Alternativamente, pueden unirse composiciones de la presente invención, o bien de manera covalente o bien de  
10 manera no covalente, a la superficie de tales vehículos portadores.

La formación y utilización de liposoma y preparaciones de tipo liposoma como posibles portadores farmacológicos las conocen generalmente los expertos en la materia (véanse por ejemplo, Lasic, Trends Biotechnol julio de 1998; 16(7):307-21; Takakura, Nippon Rinsho marzo de 1998; 56(3):691-5; Chandran *et al.*, Indian J Exp Biol. agosto de 1997; 35(8):801-9; Margalit, Crit Rev Ther Drug Carrier Syst. 1995;12(2-3):233-61; patente US 5.567.434; patente US 5.552.157; patente US 5.565.213; patente US 5.738.868 y patente US 5.795.587).

Se han utilizado satisfactoriamente liposomas con varios tipos de células que normalmente son difíciles de transfectar mediante otros procedimientos, incluyendo suspensiones de células T, cultivos de hepatocitos primarios y células PC 12 (Renneisen *et al.*, J Biol Chem. 25 de septiembre de 1990; 265(27):16337-42; Muller *et al.*, DNA Cell Biol. abril de 1990; 9(3):221-9). Además, los liposomas están libres de las limitaciones de longitud del ADN que son típicas de sistemas de administración basados en virus. Se han utilizado eficazmente liposomas para introducir genes, diversos fármacos, agentes radioterápicos, enzimas, virus, factores de transcripción, efectores alostéricos y similares, en una variedad de líneas celulares cultivadas y animales. Además, la utilización de liposomas no parece estar asociada con respuestas autoinmunitarias o toxicidad inaceptable tras la administración sistémica.

En determinadas realizaciones, se forman liposomas a partir de fosfolípidos que se dispersan en un medio acuoso y forman espontáneamente vesículas bicapa concéntricas multilaminares (también denominadas vesículas multilaminares (MLV)).

Alternativamente, en otras realizaciones, la invención proporciona formulaciones de nanocápsulas farmacéuticamente aceptables de las composiciones de la presente invención. Las nanocápsulas pueden atrapar generalmente compuestos de una manera estable y reproducible (véase, por ejemplo, Quintanar-Guerrero *et al.*, Drug Dev Ind Pharm. diciembre de 1998; 24(12):1113-28). Para evitar efectos secundarios debidos a sobrecarga polimérica intracelular, tales partículas ultrafinas (tamaño de aproximadamente 0,1  $\mu\text{m}$ ) pueden diseñarse utilizando polímeros que pueden degradarse *in vivo*. Tales partículas pueden prepararse tal como se describe, por ejemplo, por Couvreur *et al.*, Crit Rev Ther Drug Carrier Syst. 1988; 5(1):1-20; zur Muhlen *et al.*, Eur J Pharm Biopharm. marzo de 1998; 45(2):149-55; Zambaux *et al.* J Controlled Release. 2 de enero de 1998; 50(1-3):31-40; y la patente US 5.145.684.

#### 40 Procedimientos terapéuticos para EII

Los enfoques inmunológicos para la terapia de EII se basan en el reconocimiento de que EII representa una respuesta inmunitaria mucosa "anómala" frente a bacterias dentro de la luz del tubo gastrointestinal. No se ha descrito la naturaleza molecular precisa del/de los antígeno(s) bacteriano(s) reconocido(s) por el sistema inmunitario.

La inmunoterapia de EII se centra generalmente en inducir respuestas inmunitarias humorales, respuestas inmunitarias celulares, o ambas, con el objetivo de inducir tolerancia frente a un antígeno de bacterias entéricas particular, conduciendo así a una disminución de la inflamación en el intestino. Además, la inducción de células cooperadoras T CD4<sup>+</sup> es necesaria con el fin de inducir o bien anticuerpos o bien células T CD8<sup>+</sup> citotóxicas. Los antígenos polipeptídicos que son selectivos o de manera ideal específicos para bacterias asociadas con EII ofrecen un poderoso enfoque para inducir respuestas inmunitarias antiinflamatorias que o bien previenen o bien mejoran una respuesta inmunitaria aberrante frente a antígenos bacterianos asociados con EII, y son un aspecto importante de la presente invención.

Por tanto, en aspectos adicionales de la presente invención, las composiciones farmacéuticas tal como se describen en la presente memoria pueden utilizarse para estimular una respuesta inmunitaria contra antígenos bacterianos asociados con EII. En una realización de la presente invención, la respuesta inmunitaria inducida comprende anticuerpos que bloquean la interacción de un antígeno bacteriano con un receptor huésped. En una realización particular, los anticuerpos inducidos por las composiciones de la presente invención bloquean la interacción entre flagelina y TLR5, mejorando así la cascada proinflamatoria iniciada por la activación de NFKB. Alternativamente, las composiciones de la presente invención inducirán anticuerpos que estimulan la sensibilidad a LPS que mejoró la hiposensibilidad en individuos con mutación del gen Nod2 asociada con EII.

65 En una realización adicional de la presente invención, una respuesta inmunitaria será de naturaleza antiinflamatoria. Por ejemplo, una respuesta inmunitaria celular en la que las células T producen citocinas antiinflamatorias. Las

citocinas antiinflamatorias pueden incluir, pero no se limitan a, IL-4, IL-5, IL-10, TGF- $\beta$ .

Dentro de tales procedimientos, las composiciones farmacéuticas tal como se describen en la presente memoria se administran a un paciente, normalmente un animal de sangre caliente, preferentemente un ser humano. Un paciente puede estar afectado o no por EII. Tal como se comentó anteriormente, la administración de las composiciones farmacéuticas puede ser mediante cualquier procedimiento adecuado, incluyendo administración por vías intravenosa, intraperitoneal, intramuscular, subcutánea, intranasal, intradérmica, anal, vaginal, tópica y oral.

Dentro de determinadas realizaciones, la inmunoterapia puede ser inmunoterapia activa, en la que el tratamiento se basa en la estimulación *in vivo* del sistema inmunitario del huésped endógeno para reaccionar contra bacterias con la administración de agentes modificadores de la respuesta inmunitaria o inmunomoduladores (tales como polipéptidos y polinucleótidos tal como se proporcionan en la presente memoria).

Dentro de otras realizaciones, la inmunoterapia puede ser inmunoterapia pasiva, en la que el tratamiento implica la administración de agentes con reactividad inmunitaria antibacteriana establecida (tal como células efectoras o anticuerpos) que puede mediar directa o indirectamente efectos antibacterianos y no depende necesariamente de un sistema inmunitario del huésped intacto. Los ejemplos de células efectoras incluyen células T tal como se comentó anteriormente, linfocitos T (tales como linfocitos T citotóxicos CD8<sup>+</sup> y linfocitos cooperadores T CD4<sup>+</sup>), linfocitos citolíticos (tales como linfocitos citolíticos naturales y linfocitos citolíticos activados por linfocinas), células B y células presentadoras de antígeno (tales como células dendríticas y macrófagos) que expresan un polipéptido proporcionado en la presente memoria. Los receptores de células T y receptores de anticuerpos específicos para los polipéptidos mencionados en la presente memoria pueden clonarse, expresarse y transferirse en otros vectores o células efectoras para la inmunoterapia adoptiva. Los polipéptidos proporcionados en la presente memoria también pueden utilizarse para generar anticuerpos o anticuerpos anti-idiotípicos (tal como se describió anteriormente y en la patente US n.º 4.918.164) para inmunoterapia pasiva.

Pueden marcarse anticuerpos monoclonales con una variedad de etiquetas para utilizaciones selectivas deseadas en detección, ensayos de diagnóstico o aplicaciones terapéuticas (tal como se describe en las patentes US n.ºs 6.090.365; 6.015.542; 5.843.398; 5.595.721 y 4.708.930). En cada caso, la unión del anticuerpo monoclonal marcado al sitio determinante del antígeno indicará la detección o administración de un agente terapéutico particular al determinante antigénico en la célula no normal. Un objeto adicional de esta invención es proporcionar el anticuerpo monoclonal específico adecuadamente marcado para lograr tales utilizaciones selectivas del mismo.

Generalmente pueden obtenerse células efectoras en cantidades suficientes para la inmunoterapia adoptiva mediante crecimiento *in vitro*, tal como se describe en la presente memoria. En la materia se conocen bien condiciones de cultivo para expandir células efectoras específicas de antígeno individuales a un número de varios billones con retención de reconocimiento de antígeno *in vivo*. Tales condiciones de cultivo *in vitro* utilizan normalmente estimulación intermitente con antígeno, con frecuencia en presencia de citocinas (tales como IL-2) y células de alimentación que no se dividen. Tal como se indicó anteriormente, pueden utilizarse polipéptidos inmunorreactivos tal como se proporcionan en la presente memoria para expandir rápidamente cultivos de células T Específicas de antígeno con el fin de generar un número suficiente de células para inmunoterapia. En particular, pueden pulsarse células presentadoras de antígeno, tales como células dendríticas, macrófagos, monocitos, fibroblastos y/o células B, con polipéptidos inmunorreactivos o transfectarse con uno o más polinucleótidos utilizando técnicas convencionales bien conocidas en la materia. Por ejemplo, pueden transfectarse células presentadoras de antígeno con un polinucleótido que presenta un promotor apropiado para aumentar la expresión en un virus recombinante u otro sistema de expresión. Las células efectoras cultivadas para su utilización en terapia deben poder crecer y distribuirse ampliamente, y sobrevivir a largo plazo *in vivo*. Estudios han mostrado que puede inducirse que crezcan células efectoras cultivadas *in vivo* y que sobrevivan a largo plazo en números sustanciales mediante estimulación repetida con antígeno complementado con IL-2 (véase, por ejemplo, Cheever *et al.*, Immunological Reviews 157:177, 1997).

Alternativamente, puede introducirse un vector que expresa un polipéptido mencionado en la presente memoria en células presentadoras de antígeno tomadas de un paciente y propagarse mediante clonación *ex vivo* para trasplantarlas de vuelta al interior del mismo paciente. Pueden volver a introducirse células transfectadas en el paciente utilizando cualquier medio conocido en la materia, preferentemente en forma estéril mediante administración intravenosa, intracavitaria, intraperitoneal.

Las vías y frecuencia de administración de las composiciones terapéuticas descritas en la presente memoria, así como la dosificación, variarán de un individuo a otro individuo, y pueden establecerse fácilmente utilizando técnicas convencionales. En general, las composiciones farmacéuticas y vacunas pueden administrarse mediante inyección (por ejemplo, intracutánea, intramuscular, intravenosa o subcutánea), por vía intranasal (por ejemplo, mediante aspiración) o por vía oral. Preferentemente, pueden administrarse entre 1 y 10 dosis a lo largo de un periodo de 52 semanas. Preferentemente, se administran 6 dosis, a intervalos de 1 mes, y posteriormente pueden administrarse vacunas de refuerzo de manera periódica. Protocolos alternativos pueden ser apropiados para pacientes individuales. Una dosis adecuada es una cantidad de un compuesto que, cuando se administra tal como se describió anteriormente, puede fomentar una respuesta inmunitaria antibacteriana apropiada, y es al menos un 10-50%

superior al nivel basal (es decir, sin tratar). Alternativamente, una dosis adecuada es una cantidad de un compuesto que, cuando se administra tal como se describió anteriormente, puede tratar la inflamación en el colon asociada con EII. Una disminución de este tipo es de al menos el 10-50% por debajo del nivel sin tratar. Tal respuesta puede monitorizarse midiendo los anticuerpos antibacterianos en un paciente o mediante generación dependiente de vacuna de células efectoras citolíticas que pueden reconocer antígenos bacterianos *in vitro* identificados a partir de bacterias aisladas del paciente. Tales vacunas deben poder provocar una respuesta inmunitaria que conduce a un desenlace clínico mejorado (por ejemplo, remisiones más frecuentes, supervivencia libre de enfermedad completa o parcial o más prolongada) en pacientes vacunados en comparación con pacientes no vacunados. En general, para composiciones farmacéuticas y vacunas que comprenden uno o más polipéptidos, la cantidad de cada polipéptido presente en una dosis oscila entre aproximadamente 25 µg y 5 mg por kg de huésped. Los tamaños de dosis adecuados variarán con el tamaño del paciente, pero oscilarán normalmente entre aproximadamente 0,1 ml y aproximadamente 5 ml.

En general, una dosificación y régimen de tratamiento apropiados proporcionan el/los principio(s) activo(s) en una cantidad suficiente para proporcionar beneficio terapéutico y/o profiláctico. Puede monitorizarse una respuesta de este tipo estableciendo un desenlace clínico mejorado (por ejemplo, una disminución de la inflamación en el intestino, disminución de diarrea, disminución de los requisitos de esteroides o requisito de otras terapias inmunosupresoras, disminución de anemia o disminución del índice de actividad de enfermedad de Crohn (CDAI)) en pacientes tratados en comparación con pacientes no tratados. Los aumentos de respuestas inmunitarias antiinflamatorias apropiadas frente a una proteína bacteriana se correlacionan generalmente con un desenlace clínico mejorado. Tales respuestas inmunitarias pueden evaluarse generalmente utilizando ensayos convencionales de proliferación, citotoxicidad o citocina, que pueden realizarse utilizando muestras obtenidas de un paciente antes y después del tratamiento.

#### Composiciones, procedimientos y kits de detección y diagnóstico

Hay una variedad de formatos de ensayo conocidos por los expertos habituales en la materia para utilizar un agente de unión para detectar marcadores polipeptídicos en una muestra. Véase, por ejemplo, Harlow y Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. En general, la presencia o ausencia de una bacteria asociada con EII en un paciente puede determinarse mediante (a) poner en contacto una muestra biológica obtenida de un paciente con un agente de unión; (b) detectar en la muestra un nivel de polipéptido que se une al agente de unión; y (c) comparar el nivel de polipéptido con un valor de corte predeterminado.

En una realización preferida, el ensayo implica la utilización de agente de unión inmovilizado sobre un soporte sólido para unirse al, y retirar el, polipéptido del resto de la muestra. Entonces puede detectarse el polipéptido unido utilizando un reactivo de detección que contiene un grupo indicador y que se une específicamente al complejo agente de unión/polipéptido. Tales reactivos de detección pueden comprender, por ejemplo, un agente de unión que se une específicamente al polipéptido o un anticuerpo u otro agente que se une específicamente al agente de unión, tal como un anticuerpo anti-inmunoglobulina, proteína G, proteína A o una lectina. Alternativamente, puede utilizarse un ensayo competitivo, en el que se marca un polipéptido con un grupo indicador y se permite que se una al agente de unión inmovilizado tras la incubación del agente de unión con la muestra. El grado en el cual los componentes de la muestra inhiben la unión del polipéptido marcado al agente de unión es indicativo de la reactividad de la muestra con el agente de unión inmovilizado. Los polipéptidos adecuados para su utilización dentro de tales ensayos incluyen proteínas bacterianas de longitud completa y partes de polipéptido de las mismas a las que se une el agente de unión, tal como se describió anteriormente.

El soporte sólido puede ser cualquier material conocido por los expertos habituales en la materia al que puede fijarse la proteína bacteriana. Por ejemplo, el soporte sólido puede ser un pocillo de prueba en una placa de microtitulación o una membrana de nitrocelulosa u otra adecuada. Alternativamente, el soporte puede ser una perla o disco, tal como vidrio, fibra de vidrio, látex o un material de plástico tal como poliestireno o poli(cloruro de vinilo). El soporte también puede ser una partícula magnética o un sensor de fibra óptica, tal como los dados a conocer, por ejemplo, en la patente US n.º 5.359.681. El agente de unión puede inmovilizarse en el soporte sólido utilizando una variedad de técnicas conocidas por los expertos en la materia, que se describen ampliamente en la bibliografía científica y de patentes. En el contexto de la presente invención, el término "inmovilización" se refiere tanto a asociación no covalente, tal como adsorción, como a fijación covalente (que puede ser un enlace directo entre el agente y grupos funcionales en el soporte o puede ser un enlace mediante un agente de reticulación). Se prefiere la inmovilización mediante adsorción a un pocillo en una placa de microtitulación o a una membrana. En tales casos, puede lograrse la adsorción poniendo en contacto el agente de unión, en un tampón adecuado, con el soporte sólido durante una cantidad de tiempo adecuada. El tiempo de contacto varía con la temperatura, pero es normalmente de entre aproximadamente 1 hora y aproximadamente 1 día. En general, poner en contacto un pocillo de una placa de microtitulación de plástico (tal como poliestireno o poli(cloruro de vinilo)) con una cantidad de agente de unión que oscila entre aproximadamente 10 ng y aproximadamente 10 µg, y preferentemente de aproximadamente 100 ng a aproximadamente 1 µg, es suficiente para inmovilizar una cantidad adecuada de agente de unión.

La fijación covalente de un agente de unión a un soporte sólido puede lograrse generalmente haciendo reaccionar en primer lugar el soporte con un reactivo bifuncional que reaccionará tanto con el soporte como con un grupo

funcional, tal como un grupo hidroxilo o amino, en el agente de unión. Por ejemplo, el agente de unión puede fijarse covalentemente a soportes que presentan un recubrimiento polimérico apropiado utilizando benzoquinona o mediante condensación de un grupo aldehído en el soporte con una amina y un hidrógeno activo en la pareja de unión (véase, por ejemplo, Pierce Immunotechnology Catalog and Handbook, 1991, en A1 2-A13).

5 En determinadas realizaciones, el ensayo es un ensayo de tipo sándwich de dos anticuerpos. Este ensayo puede realizarse poniendo en primer lugar en contacto un anticuerpo que se ha inmovilizado sobre un soporte sólido, comúnmente el pocillo de una placa de microtitulación, con la muestra, de tal manera que se permite que se unan los polipéptidos dentro de la muestra al anticuerpo inmovilizado. Entonces se elimina la muestra no unida de los complejos polipéptido-anticuerpo inmovilizados y se añade un reactivo de detección (preferentemente un segundo anticuerpo que puede unirse a un sitio diferente en el polipéptido) que contiene un grupo indicador. Entonces se determina la cantidad de reactivo de detección que permanece unida al soporte sólido utilizando un procedimiento apropiado para el grupo indicador específico.

15 Más específicamente, una vez inmovilizado el anticuerpo en el soporte tal como se describió anteriormente, normalmente se bloquean los sitios de unión a proteína restantes en el soporte. Cualquier agente de bloqueo adecuado conocido por los expertos habituales en la materia, tal como albúmina sérica bovina o Tween 20™ (Sigma Chemical Co., St. Louis, MO). Entonces se incuba el anticuerpo inmovilizado con la muestra, y se deja que el polipéptido se una al anticuerpo. La muestra puede diluirse con un diluyente adecuado, tal como solución salina tamponada con fosfato (PBS) antes de la incubación. En general, un tiempo de contacto apropiado (es decir, tiempo de incubación) es un periodo de tiempo que es suficiente para detectar la presencia de un polipéptido bacteriano asociado con EII dentro de una muestra obtenida de un individuo con EII de al menos aproximadamente el 95% de la lograda en el equilibrio entre polipéptido unido y no unido. Los expertos habituales en la materia reconocerán que el tiempo necesario para alcanzar el equilibrio puede determinarse fácilmente sometiendo a ensayo el nivel de unión que se produce a lo largo de un periodo de tiempo. A temperatura ambiente, un tiempo de incubación de aproximadamente 30 minutos es generalmente suficiente.

25 Entonces puede eliminarse la muestra no unida lavando el soporte sólido con un tampón apropiado, tal como PBS que contiene el 0,1% de Tween 20™. Entonces puede añadirse el segundo anticuerpo, que contiene un grupo indicador, al soporte sólido. Los grupos indicadores preferidos incluyen los grupos mencionados anteriormente.

30 Entonces se incuba el reactivo de detección con el complejo anticuerpo-polipéptido inmovilizado durante una cantidad de tiempo suficiente para detectar el polipéptido unido. Una cantidad de tiempo apropiada puede determinarse generalmente sometiendo a ensayo el nivel de unión que se produce a lo largo de un periodo de tiempo. Entonces se elimina reactivo de detección no unido y se detecta reactivo de detección unido utilizando el grupo indicador. El procedimiento empleado para detectar el grupo indicador depende de la naturaleza del grupo indicador. Para grupos radiactivos, generalmente son apropiados procedimientos de recuento por centelleo o autorradiográficos. Pueden utilizarse procedimientos espectroscópicos para detectar tintes, grupos luminiscentes y grupos fluorescentes. La biotina puede detectarse utilizando avidina, acoplada a un grupo indicador diferente (comúnmente un grupo radiactivo o fluorescente o una enzima). Generalmente pueden detectarse grupos indicadores enzimáticos mediante la adición de sustrato (generalmente durante un periodo de tiempo específico), seguido por análisis espectroscópico u otro de los productos de reacción.

35 Para determinar la presencia o ausencia de EII, generalmente se compara la señal detectada a partir del grupo indicador que permanece unido al soporte sólido con una señal que corresponde a un valor de corte predeterminado. En una realización preferida, el valor de corte para la detección de un antígeno bacteriano asociado con EII es la señal media promedio obtenida cuando se incuba el anticuerpo inmovilizado con muestras de pacientes sin EII. En general, una muestra que genera una señal que está tres desviaciones estándar por encima del valor de corte predeterminado se considera positiva para antígenos bacterianos asociados con EII. En una realización preferida alternativa, el valor de corte se determina utilizando una curva de receptor operador, según el procedimiento de Sackett *et al.*, Clinical Epidemiology: A Basic Science for Clinical Medicine, Little Brown and Co., 1985, págs. 106-7. En resumen, en esta realización, puede determinarse el valor de corte a partir de un diagrama de pares de tasas de verdaderos positivos (es decir, sensibilidad) y tasas de falsos positivos (100%-especificidad) que corresponden a cada posible valor de corte para el resultado de la prueba de diagnóstico. El valor de corte en el diagrama que está más cerca de la esquina superior izquierda (es decir, el valor más que encierra el mayor área) es el valor de corte más preciso, y una muestra que genera una señal que es superior a la del valor de corte determinado mediante este procedimiento puede considerarse positiva. Alternativamente, el valor de corte puede desplazarse a la izquierda a lo largo del diagrama para minimizar la tasa de falsos positivos, o a la derecha para minimizar la tasa de falsos negativos. En general, una muestra que genera una señal que es superior al valor de corte determinado mediante este procedimiento se considera positiva para una bacteria asociada con EII.

45 En una realización relacionada, el ensayo se realiza en un formato de flujo a través o de prueba con tira, en el que el agente de unión se inmoviliza sobre una membrana, tal como nitrocelulosa. En la prueba de flujo a través, los polipéptidos dentro de la muestra se unen al agente de unión inmovilizado a medida que la muestra pasa a través de la membrana. Entonces se une un segundo agente de unión marcado al complejo agente de unión-polipéptido a medida que fluye una disolución que contiene el segundo agente de unión a través de la membrana. Entonces

puede realizarse la detección del segundo agente de unión unido tal como se describió anteriormente. En el formato de prueba con tira, se sumerge un extremo de la membrana a la que está unido el agente de unión en una disolución que contiene la muestra. La muestra migra a lo largo de la membrana a través de una región que contiene un segundo agente de unión y a la zona de agente de unión inmovilizado. La concentración de segundo agente de unión en la zona de anticuerpo inmovilizado indica la presencia de un antígeno bacteriano asociado con EII. Normalmente, la concentración de segundo agente de unión en ese sitio genera un patrón, tal como una línea, que puede leerse visualmente. La ausencia de tal patrón indica un resultado negativo. En general, la cantidad de agente de unión inmovilizado sobre la membrana se selecciona para generar un patrón visualmente apreciable cuando la muestra biológica contiene un nivel de polipéptido que sería suficiente para generar una señal positiva en el ensayo de tipo sándwich de dos anticuerpos, en el formato comentado anteriormente. Agentes de unión preferidos para su utilización en tales ensayos son anticuerpos y fragmentos de unión a antígeno de los mismos. Preferentemente, la cantidad de anticuerpo inmovilizado sobre la membrana oscila entre aproximadamente 25 ng y aproximadamente 1 µg, y más preferentemente desde aproximadamente 50 ng hasta aproximadamente 500 ng. Tales pruebas pueden realizarse normalmente con una cantidad muy pequeña de muestra biológica.

Evidentemente, existen muchos otros protocolos de ensayo que son adecuados para su utilización con las proteínas bacterianas o agentes de unión de la presente invención. Se pretende que las descripciones anteriores sean únicamente a modo de ejemplo. Por ejemplo, resultará evidente para los expertos habituales en la materia que los protocolos anteriores pueden modificarse fácilmente para utilizar polipéptidos bacterianos para detectar anticuerpos que se unen a tales polipéptidos en una muestra biológica. La detección de tales anticuerpos específicos para proteínas bacterianas puede correlacionarse con la presencia de un antígeno asociado con EII.

Además, o alternativamente, también pueden detectarse bacterias asociadas con EII basándose en la presencia de células T que reaccionan específicamente con una proteína bacteriana en una muestra biológica. Dentro de determinados procedimientos, se incuba una muestra biológica que comprende células T CD4<sup>+</sup> y/o CD8<sup>+</sup> aisladas de un paciente con un polipéptido bacteriano, un polinucleótido que codifica para tal polipéptido y/o una CPA que expresa al menos una parte inmunogénica de tal polipéptido, y se detecta la presencia o ausencia de activación específica de las células T. Las muestras biológicas adecuadas incluyen, pero no se limitan a, células T aisladas. Por ejemplo, pueden aislarse células T de un paciente mediante técnicas rutinarias (tales como mediante centrifugación por gradiente de densidad de Ficoll/Hypaque de linfocitos de sangre periférica). En una realización particular de la presente invención, pueden aislarse células T a partir de muestras de linfocitos intraepiteliales (LIE) o linfocitos de la lámina propia (LLP) que se originan de biopsias de colon. Pueden incubarse células T *in vitro* durante 2-9 días (normalmente 4 días) a 37°C con polipéptido (por ejemplo, 5 - 25 µg/ml). Puede ser deseable incubar otra alícuota de una muestra de células T en ausencia de polipéptido bacteriano para servir como control. Para células T CD4<sup>+</sup>, la activación se detecta preferentemente evaluando la proliferación de las células T. Para células T CD8<sup>+</sup>, la activación se detecta preferentemente evaluando la actividad citolítica. Un nivel de proliferación que es al menos dos veces superior y/o un nivel de actividad citolítica que es al menos un 20% superior al de pacientes libres de enfermedad indica la presencia de bacterias asociadas con EII en el paciente.

Tal como se indicó anteriormente, también, o alternativamente, puede detectarse EII basándose en el nivel de ARNm que codifica para una proteína bacteriana en una muestra biológica. Por ejemplo, pueden emplearse al menos dos cebadores oligonucleotídicos en un ensayo basado en reacción en cadena de la polimerasa (PCR) para amplificar una parte de un ADNc bacteriano derivado de una muestra biológica, en el que al menos uno de los cebadores oligonucleotídicos es específico para (es decir, se hibrida con) un polinucleótido que codifica para la proteína bacteriana. Entonces se separa el ADNc amplificado y se detecta utilizando técnicas bien conocidas en la materia, tales como electroforesis en gel.

De manera similar, pueden utilizarse sondas oligonucleotídicas que se hibridan específicamente con un polinucleótido que codifica para una proteína bacteriana en un ensayo de hibridación para detectar la presencia de polinucleótido que codifica para la proteína bacteriana en una muestra biológica.

Para permitir la hibridación en condiciones de ensayo, las sondas y los cebadores oligonucleotídicos deben comprender una secuencia de oligonucleótido que presente al menos aproximadamente el 60%, preferentemente al menos aproximadamente el 75% y más preferentemente al menos aproximadamente el 90% de identidad con una parte de un polinucleótido que codifica para una proteína bacteriana de la invención que presenta menos 10 nucleótidos, y preferentemente al menos 20 nucleótidos de longitud. Preferentemente, las sondas y/o los cebadores oligonucleotídicos se hibridan con un polinucleótido que codifica para un polipéptido descrito en la presente memoria en condiciones de rigurosidad moderada, tal como se definió anteriormente. Las sondas y/o los cebadores oligonucleotídicos que pueden emplearse satisfactoriamente en los procedimientos de diagnóstico descritos en la presente memoria presentan preferentemente al menos 10-40 nucleótidos de longitud. Los cebadores oligonucleotídicos pueden comprender al menos 10 nucleótidos contiguos, más preferentemente al menos 15 nucleótidos contiguos, de una molécula de ADN que presenta una secuencia tal como se da a conocer en la presente memoria. En la materia se conocen bien técnicas tanto para ensayos basados en PCR como para ensayos de hibridación (véanse, por ejemplo, Mullis *et al.*, Cold Spring Harbor Symp. Quart. Biol., 51:263, 1987; Erlich ed., PCR Technology, Stockton Press, NY, 1989).

- 5 Un ensayo preferido emplea RT-PCR, en el que se aplica PCR junto con transcripción inversa. Normalmente, se extrae ARN de una muestra biológica, tal como tejido de biopsia, y se somete a transcripción inversa para producir moléculas de ADNc. La amplificación por PCR utilizando al menos un cebador específico genera una molécula de ADNc, que puede separarse y visualizarse utilizando, por ejemplo, electroforesis en gel. La amplificación puede realizarse con muestras biológicas tomadas de un paciente de prueba y de un individuo que no estaba afectado por un cáncer. La reacción de amplificación puede realizarse en varias diluciones de ADNc que abarcan dos órdenes de magnitud. Un aumento de dos veces o más en la expresión en varias diluciones de la muestra de paciente de prueba en comparación con las mismas diluciones de la muestra no cancerosa se considera normalmente positivo.
- 10 En otra realización, pueden utilizarse las composiciones descritas en la presente memoria como marcadores para la progresión de EII. En esta realización, pueden realizarse ensayos tal como se describió anteriormente para el diagnóstico de EII a lo largo del tiempo, y evaluarse el cambio en el nivel de polipéptido(s) o polinucleótido(s) reactivo(s). Por ejemplo, los ensayos pueden realizarse cada 24-72 horas durante un periodo de 6 meses a 1 año, y posteriormente realizarse según se necesite. En general, la EII está en progresión en aquellos pacientes en los que el nivel de polipéptido o polinucleótido detectado aumenta a lo largo del tiempo. En cambio, la EII no está en progresión cuando el nivel de polipéptido o polinucleótido reactivo o bien permanece constante o bien disminuye con el tiempo.
- 15 En una realización adicional, pueden utilizarse las composiciones descritas en la presente memoria para monitorizar el nivel de anticuerpos específicos para, y/o sensibilidad de células T frente a, una proteína bacteriana asociada con EII como medida de la progresión de EII. En general, la EII está en progresión en aquellos pacientes en los que el nivel de anticuerpos que se unen a un polipéptido o codificados por un polinucleótido descrito en la presente memoria, que se detecta aumenta a lo largo del tiempo. En cambio, la EII no está en progresión cuando el nivel de anticuerpos reactivos permanece constante o disminuye a lo largo del tiempo.
- 20 Pueden realizarse determinados ensayos de diagnóstico *in vivo* directamente en una lesión en el colon. Un ensayo de este tipo implica poner en contacto células de una lesión con un agente de unión. Entonces puede detectarse e agente de unión unido directa o indirectamente mediante un grupo indicador. Tales agentes de unión también pueden utilizarse en aplicaciones histológicas. Alternativamente, pueden utilizarse sondas polinucleotídicas dentro de tales aplicaciones.
- 25 Tal como se indicó anteriormente, para mejorar la sensibilidad, pueden someterse a ensayo múltiples proteínas bacterianas dentro de una muestra dada. Resultará evidente que pueden combinarse agentes de unión específicos para diferentes proteínas, anticuerpos, o células T específicas para los mismos proporcionados en la presente memoria dentro de un único ensayo. Además, pueden utilizarse simultáneamente múltiples cebadores o sondas. La selección de proteínas bacterianas puede basarse en experimentos de rutina para determinar combinaciones que dan como resultado sensibilidad óptima. Además, o alternativamente, pueden combinarse ensayos para detectar proteínas bacterianas, anticuerpos, o células T específicas para los mismos, proporcionados en la presente memoria con ensayos para detectar otros antígenos bacterianos conocidos o marcadores genéticos tales como la mutación NOD2.
- 30 La presente invención proporciona además kits para su utilización dentro de cualquiera de los procedimientos de diagnóstico anteriores. Tales kits comprenden normalmente dos o más componentes necesarios para realizar un ensayo de diagnóstico. Los componentes pueden ser compuestos, reactivos, recipientes y/o equipo. Por ejemplo, un recipiente dentro de un kit puede contener un anticuerpo monoclonal o fragmento del mismo que se une específicamente a una proteína bacteriana. Tales anticuerpos o fragmentos pueden proporcionarse unidos a un material de soporte, tal como se describió anteriormente. Uno o más recipientes adicionales pueden contener elementos, tales como reactivos o tampones, que van a utilizarse en el ensayo. Además, o alternativamente, tales kits pueden contener un reactivo de detección tal como se describió anteriormente que contiene un grupo indicador adecuado para la detección directa o indirecta de unión de anticuerpo.
- 35 Alternativamente, puede diseñarse un kit para detectar el nivel de ARNm que codifica para una proteína bacteriana en una muestra biológica. Tales kits comprenden generalmente al menos un cebador o sonda oligonucleotídica, tal como se describió anteriormente, que se hibrida con un polinucleótido que codifica para una proteína bacteriana. Un oligonucleótido de este tipo puede utilizarse, por ejemplo, dentro de un ensayo de PCR o de hibridación. Los componentes adicionales que pueden estar presentes dentro de tales kits incluyen un segundo oligonucleótido y/o un reactivo de diagnóstico o recipiente para facilitar la detección de un polinucleótido que codifica para una proteína bacteriana.
- 40 En una realización alternativa, puede diseñarse un kit para detectar el nivel de anticuerpos específicos para una proteína bacteriana asociada con EII en una muestra biológica.

Los siguientes ejemplos se ofrecen a modo de ilustración y no a modo de limitación.

**Ejemplos**

**Ejemplo 1**

5 IDENTIFICACIÓN DE ANTÍGENOS BACTERIANOS ASOCIADOS CON EII A PARTIR DE UNA BIBLIOTECA DE EXPRESIÓN DE CIZALLADURA AL AZAR GENÓMICA DE BACTERIAS CECALES DE RATÓN

Se construyó una biblioteca de expresión de cizalladura al azar genómica de bacterias cecales de ratón sonicando ADN genómico de bacterias cecales de ratón C3H/HeJ para producir tamaños de fragmento de aproximadamente 0,1 a 5,0 kpb. Se trataron 14 µg de ADN sonicado con ADN polimerasa I, fragmento de Klenow, durante 30 minutos seguido por polimerasa de Pfu durante 30 minutos para producir fragmentos con extremos romos. Entonces se ligaron adaptadores de EcoRI a los fragmentos y después se fosforilaron los adaptadores con polinucleótido cinasa de *E. coli*. A continuación se fraccionaron los fragmentos con una columna Sephacryl S400 y finalmente se ligaron a un vector Lambda ZAP Express (Stratagene). Entonces se empaquetó vector ligado con extracto de empaquetamiento Gigapack III Gold (Stratagene) y se sembró en placas la biblioteca no amplificada con células huésped *E. coli* XL-1 Blue MRF' en placas de LB agarosa a una concentración de 25.000 unidades formadoras de placa (upf) por placa para 15 placas. Tras la incubación a 42°C durante 4 horas, se añadieron filtros de nitrocelulosa, empapados en IPTG 10 mM, y se incubaron las placas a 37°C durante la noche. Se retiraron los filtros y se lavaron 3X con PBS que contenía el 0,1% de Tween 20 (PBST), se bloquearon durante 1 hora con el 1% de BSA en PBST, se lavaron 3X con PBST y después se incubaron durante la noche a 4°C en suero, preadsorbido con proteínas de *E. coli*, de ratones C3H/HeJ Bir con enfermedad inflamatoria del intestino. Tras lavar 3X con PBST, se incubaron los filtros con anticuerpo secundario de cabra anti-IgG, IgA, IgM de ratón conjugado con fosfatasa alcalina durante 1 hora a temperatura ambiente. Finalmente se lavaron los filtros 3X con PBST, 2X con tampón de fosfatasa alcalina y se revelaron con BCIP/NBT. Se purificaron los clones positivos utilizando la misma técnica; se escindió el fagémido; y se secuenció el ADN de plásmido resultante y se hizo una búsqueda del mismo en las bases de datos de GenBank. Las secuencias que mostraron algún grado de similitud con secuencias conocidas en la base de datos se indican en la tabla 2. Las secuencias que no mostraron similitud significativa con secuencias conocidas en la base de datos se muestran en la tabla 3.

**TABLA 2**

SECUENCIAS BACTERIANAS QUE MOSTRARON ALGÚN GRADO DE SIMILITUD CON SECUENCIAS CONOCIDAS EN LA BASE DE DATOS

SEC ID de ADNc	SEC ID de PRO	Nombre del clon	ID del clon	kpb de inserto	Blastn	Blastx
1	38	Cbir-1	76779	1,3	Sin coincidencia	(73%) proteína flagelina A / (48%) FlaB [ <i>Butyrivibrio fibrisolvens</i> ] (AF026812)
2	39	Cbir-2	76780	0,4	Sin coincidencia	(73%) 3-isopropilmalato deshidrogenasa VC2491- <i>Vibrio cholerae</i> (cepa N16961 de grupo O1)
3	40, 41	Cbir-3	76959	0,8	Sin coincidencia	(53%) proteína de movilidad A- <i>Termotoga maritima</i> (cepa MSB8)
5		Cbir-5	76961	1,2	Sin coincidencia	(44%) proteína de quimiotaxis aceptadora de metilo [ <i>Bacillus halodurans</i> ] (AP001520)
6	42	Cbir-6	76781	1,0	Sin coincidencia	(56%) proteína ribosómica L6 (BL10)
7		Cbir-8	76962	3,0	Sin coincidencia	(60%) cadena grande de acetohidroxiácido sintasa - <i>Brevibacterium flavum</i>
10		Cbir-12	76964	2,0	(82%) Igual que Blastx	(58%) fosfoenolpiruvato carboxinasa VC2738- <i>Vibrio cholerae</i> (cepa N16961 de grupo O1)
12		Cbir-14	76965	1,7	Sin coincidencia	(49%) regulador de huecos [ <i>Staphylococcus aureus</i> ] (AJ133520)
13		Cbir-15	76966	1,0	(82%) Igual que Blastx	(72%) proteína flagelina A / FlaB ( <i>Butyrivibrio fibrisolvens</i> ) (AF026812)
14	44	Cbir-16	76967	1,2	Sin coincidencia	(58%) proteína ribosómica L6 (BL10)
15		Cbir-18	76968	0,8	Sin coincidencia	(69%) flagelina - <i>Roseburia cecicola</i> (65%) proteína flagelina A / (63%) FlaB [ <i>Butyrivibrio fibrisolvens</i> ] (AF026812)
16		Cbir-19	77530	1,6	Sin coincidencia	(51%) proteína de transporte tipo ABC - <i>Synechocystis sp.</i> (cepa PCC6803)
17	45	Cbir-20	76969	2,0	Sin coincidencia	(68%) FlaB [ <i>Butyrivibrio fibrisolvens</i> ] (AF026812)
18		Cbir-23	76970	0,8	(80%) Igual que Blastx	(84%) factor de elongación - Tu [ <i>Porphyromonas gingivalis</i> ] (AB035462)

ES 2 405 405 T3

SEC ID de ADNc	SEC ID de PRO	Nombre del clon	ID del clon	kpb de inserto	Blastn	Blastx
23	47	Cbir-32	76974	65 pb	Sin coincidencia	probablemente amidohidrolasa <i>Campylobacter jejuni</i> (cepa NCTC 11168)
24		Cbir-36	77074	1,6	(87%) genoma completo de <i>Bacillus subtilis</i>	(83%) factor de elongación TU-1 ( <i>Streptomyces ramocissimus</i> )
26	48	Cbir-39	76975	0,6	Sin coincidencia	(70%) <i>Thermotoga maritima</i> (cepa MSB8) (55%) proteína flagelina A / (56%) FlaB [ <i>Butyrivibrio fibrisolvens</i> ] (AF026812)
27		Cbir-40	77075	0,7	Sin coincidencia	(54%) flagelina - <i>Thermotoga maritima</i> (cepa MSB8) / (58%) flagelina ( <i>Bacillus halodurans</i> ) (AP001512)
29	49	Cbir-44	76977	1,0	Sin coincidencia	(45%) flagelina - <i>Thermotoga maritima</i> (cepa MSB8) / (57%) flagelina ( <i>Bacillus halodurans</i> ) (AP001512)
31		Cbir-46	77533	2,2	Sin coincidencia	(AL512975)
32		Cbir-49	77534	0,9	Sin coincidencia	(59%) flagelina - <i>Roseburia cecicola</i> (56%) proteína flagelina A / (58%) FlaB ( <i>Butyrivibrio fibrisolvens</i> ) (AF026812)
35		Cbir-62	77536	0,6	Sin coincidencia	precursor de proteína de secreción de hemolisina - <i>Helicobacter pylori</i> (cepa 26695)/ proteína de quimiotaxis aceptadora de metilo ( <i>Helicobacter pylori</i> )
36		Cbir-73	77538	2,0	Sin coincidencia	factor de elongación - TS
37		Cbir-78	77539	1,0	Sin coincidencia	(62%) flagelina - <i>Roseburia cecicola</i> (M20983) (57%) proteína flagelina A / (59%) FlaB ( <i>Butyrivibrio fibrisolvens</i> )
51		CB1-T2	73261	150	Sin coincidencia	36% arabinosidasa de <i>Streptomyces</i> secretado
54		CB1-T5	73264	1050	Sin coincidencia	33% ClpC proteasa de <i>Arabidopsis</i>
55		CB1-T7	73266	650	Sin coincidencia	54% acetiltransferasa de <i>S. cerevisiae</i>
56		CB1-T8	73267	350	Sin coincidencia	28% desconocido de <i>Bacillus subtilis</i>
58		CB1-T10	73269	150	Sin coincidencia	? 40% transportadora ABC de <i>Deinococcus</i>
59		CB1-T11	73270	1700	Sin coincidencia	52% fructosa-6-p 1-fosfotransferasa de <i>Borrelia burgdorferi</i>
60		CB1-T13	73272	250	Sin coincidencia	56% alcohol deshidrogenasa <i>Vibrio/Clostridium</i>
61		CB1-T14	73273	260	Sin coincidencia	? 53% Ag de superficie BspA de <i>Bacteroides</i>
62		CB 1-T15	73274	1000	Sin coincidencia	61% protil ARNt sintetasa de <i>Clostridium</i>
64		CB3-T1	75037	1,7	? 88% de <i>Streptomyces</i>	65% proteína OPPF de unión a ATP de transporte de oligopéptido de <i>Bacillus subtilis</i>
67		CB3-T4	75040	?>600	Sin coincidencia	36% proteína de membrana de 14,8 kDa hipotética de <i>E.coli/Staph</i>
70		CB3-T9	75044	300	Sin coincidencia	91% 2,3,4,5-tetrahidropiridina-2-carboxilato n-succiniltransferasa de <i>Helicobacter pylori</i>
72		CB3-T12	75046	1700	Sin coincidencia	52% proteína conservada de <i>Thermotoga/ Archaeoglobus / Halobacterium</i>
74		CB3-T14	75048	1300	Sin coincidencia	29% choque térmico de <i>Streptomyces</i>

TABLA 3

SECUENCIAS BACTERIANAS QUE NO MOSTRARON SIMILITUD SIGNIFICATIVA CON SECUENCIAS CONOCIDAS EN LA BASE DE DATOS

5

SEC ID de ADNc	SEC ID de PRO	Nombre del clon	ID del clon	kpb de inserto	Blastn	Blastx
4		Cbir-4	76960	2,1	Sin coincidencia	Sin coincidencia
8	43	Cbir-9	76782	0,6	Sin coincidencia	Sin coincidencia
9		Cbir-11	76963	1,5	Sin coincidencia	Sin coincidencia
11		Cbir-13	77529	1,5	Sin coincidencia	Sin coincidencia
19		Cbir-24	76971	2,3	Sin coincidencia	Sin coincidencia
20		Cbir-26	77073	1,2	Sin coincidencia	Sin coincidencia
21	46	Cbir-27	76972	0,5	Sin coincidencia	Sin coincidencia

SEC ID de ADNc	SEC ID de PRO	Nombre del clon	ID del clon	kpb de inserto	Blastn	Blastx
22		Cbir-30	76973	0,7	Sin coincidencia	Sin coincidencia
25		Cbir-37	77531	2,5	Sin coincidencia	Sin coincidencia
28		Cbir-41	76976	1,2	Sin coincidencia	Sin coincidencia
30		Cbir-45	77532	1,2	Sin coincidencia	Sin coincidencia
33		Cbir-50	77535	0,7	Sin coincidencia	Sin coincidencia
34	50	Cbir-61	77076	0,5	Sin coincidencia	Sin coincidencia
52		CB1-T3	73262	600	Sin coincidencia	Sin coincidencia
53		CB1-T4	73263	250	Sin coincidencia	Sin coincidencia
57		CB1-T9	73268	250	Sin coincidencia	Sin coincidencia
63		CB1-T16	73275	400	Sin coincidencia	Sin coincidencia
65		CB3-T2	75038	500	Sin coincidencia	Sin coincidencia
66		CB3-T3	75039	500	Sin coincidencia	Sin coincidencia
68		CB3-T5	75041	400	Sin coincidencia	Sin coincidencia
69		CB3-T6	75042	600	Sin coincidencia	Sin coincidencia
71		CB3-T10	75045	?>500	Sin coincidencia	Sin coincidencia
73		CB3-T13	75047	700	Sin coincidencia	Sin coincidencia

### Ejemplo 2

#### CLONACIÓN Y EXPRESIÓN DE FLAGELINA X

5 Se clonó una flagelina novedosa mediante amplificación por PCR a partir de ADN genómico de bacterias cecales total obtenido de ratones C3H/HeJ Bir. Se desarrollaron cebadores oligonucleotídicos con secuencia del extremo amino-terminal del clon Cbir-1 (SEC ID Nº: 1) y con una secuencia consenso derivada del extremo carboxilo-terminal de tres secuencias de flagelina que estaban lo más estrechamente relacionadas con la secuencia de Cbir-1. Se  
10 añadieron una etiqueta de seis histidinas (SEC ID Nº: 87) y un sitio de escisión de endonucleasa Nde I al cebador amino para la clonación y para la purificación de proteína recombinante y se añadió un sitio de clonación Hind III al cebador carboxilo-terminal. Se expresó la secuencia de ADNc de flagelina X de longitud completa (SEC ID Nº: 75) como constructo pET17b (Novagen, Madison, WI) y se purificó la proteína recombinante resultante con una columna de afinidad Ni-NTA (Qiagen, Valencia, CA). La secuencia de aminoácidos de flagelina X de longitud completa se  
15 representa en SEC ID Nº: 79. Tanto el análisis de inmunotransferencia de tipo Western como ensayos ELISA demostraron que esta proteína de flagelina recombinante era reactiva con anticuerpo procedente de ratones con EII.

Se construyeron tres formas truncadas de flagelina X y se utilizaron en estudios de expresión. Se clonaron los siguientes truncamientos utilizando cebadores de PCR desarrollados a partir de la secuencia de flagelina X:

- 20 i. el extremo conservado amino-terminal de la molécula (secuencia de ADNc: SEC ID Nº: 76, secuencia de aminoácidos: SEC ID Nº: 80)
- 25 ii. el extremo conservado amino-terminal más la parte variable de la molécula (secuencia de ADNc: SEC ID Nº: 77, secuencia de aminoácidos: SEC ID Nº: 81)
- iii. y la parte conservada carboxilo-terminal (secuencia de ADNc: SEC ID Nº: 78, secuencia de aminoácidos: SEC ID Nº: 82)

30 Se añadieron una etiqueta de seis histidinas (SEC ID Nº: 87) y sitios de clonación a todos los constructos para la clonación y expresión con el vector pET17b. Se expresaron los constructos en *E. coli* y se analizó la proteína recombinante mediante inmunotransferencia de tipo Western y ELISA. Ensayos de inmunotransferencia de tipo Western y ELISA con proteína recombinante amino-terminal y carboxilo-terminal demostraron que la mayor reactividad de anticuerpo era frente al extremo amino-terminal conservado. Estas proteínas recombinantes  
35 presentan utilidad en el diagnóstico, la terapia y el desarrollo de vacunas para EII.

### Ejemplo 3

#### IDENTIFICACIÓN DE LA SECUENCIA DE POLIPÉPTIDO Y DE NUCLEÓTIDOS DE LONGITUD COMPLETA DE FLAGELINA B DE *HELICOBACTER BILIS*

40 Este ejemplo describe la identificación de la secuencia de polipéptido y de nucleótidos de flagelina B de *H. bilis* de longitud completa. Las secuencias de proteína y de ADN recombinantes presentan utilidad, por ejemplo, en el diagnóstico, la terapia y el desarrollo de vacunas para EII.

45 Se ha notificado que la infección con *Helicobacter spp*, incluyendo *H. bilis*, provoca EII en ratones inmunodeficientes. Por tanto las proteínas de *Helicobacter* específicas y de *Helicobacter spp* son herramientas útiles para investigar EII

inducida por microbios. También se han implicado *Helicobacter spp*, especialmente *H. pylori*, en EII en seres humanos. Tal como se muestra en los ejemplos 1 y 2 en la presente memoria, se cree que la flagelina bacteriana está asociada con EII en el modelo de ratón. En la presente memoria se describen secuencias de nucleótidos y de polipéptido para flagelina B de *H. bilis* (FlaB). Se amplificó esta secuencia utilizando técnicas convencionales de ADN genómico total de *H. bilis* utilizando cebadores oligonucleotídicos derivados de la secuencia de flagelina B de *H. mustelae*. La secuencia de nucleótidos de longitud completa (secuencia codificante) de la flagelina B de *H. bilis* se expone en SEC ID N°: 83. La secuencia de aminoácidos codificada por este nucleótido se expone en SEC ID N°: 84. Este polipéptido mostró homología con proteínas de flagelina B de *H. mustelae* y *H. pylori*. La región conservada amino-terminal de la proteína de flagelina B de *H. bilis* incluye los residuos de aminoácido de 1 a 154 de SEC ID N°: 84, y los nucleótidos correspondientes que codifican para esta región de SEC ID N°: 83. La parte conservada carboxilo-terminal incluye los residuos de aminoácido de 365 a 514 de SEC ID N°: 84, y los nucleótidos correspondientes que codifican para esta región de SEC ID N°: 83. En resumen, estas secuencias representan dianas terapéuticas y de diagnóstico atractivas para EII.

#### 15 Ejemplo 4

##### IDENTIFICACIÓN DE LA SECUENCIA DE PROTEÍNA Y DE NUCLEÓTIDOS DE LONGITUD COMPLETA DE FLAGELINA CBIR-1

20 Este ejemplo describe la secuenciación de longitud completa del clon de flagelina Cbir-1 (secuencia parcial expuesta en SEC ID N°: 1). Este clon se obtuvo originalmente mediante examen de expresión serológica de una biblioteca de bacterias cecales de ratón con suero de ratones con EII (véase el ejemplo 1 y la tabla 2). Las secuencias de proteína y de ADN recombinantes presentan utilidad, por ejemplo, en el diagnóstico, la terapia y el desarrollo de vacunas para EII.

25 El clon Cbir-1 es una secuencia de flagelina bacteriana parcial (extremo amino más región variable) que se mostró que era altamente reactiva con el suero de ratón utilizado para examinar la biblioteca bacteriana. La proteína recombinante derivada de este clon y un recombinante que representaba el extremo amino-terminal de esta secuencia clonada también fueron altamente reactivos con suero humano y de ratón enfermo (IgG2a) y también contienen posiblemente un epítipo de células T. Estos datos indican que esta proteína es inmunogénica en el ratón y en el ser humano. En la presente memoria se describen las secuencias de nucleótidos (SEC ID N°: 85) y de proteína (SEC ID N°: 86) de flagelina CBir-1 de longitud completa. La secuencia de nucleótidos se obtuvo mediante técnicas de amplificación por PCR convencionales a partir de ADN de bacterias cecales de ratón genómico total con un cebador específico para la secuencia del clon CBir-1 y un segundo cebador derivado del extremo carboxilo conservado de secuencias de flagelina relacionadas. Alineaciones de polipéptidos confirmaron que la proteína estaba relacionada con otras proteínas de flagelina. La región conservada amino-terminal de la proteína CBir-1 incluye los residuos de aminoácido de 1 a 147 de SEC ID N°: 86, y los nucleótidos correspondientes que codifican para esta región de SEC ID N°: 85. El extremo conservado amino-terminal más la parte variable de la molécula incluye los residuos de aminoácido de 1 a 418 de SEC ID N°: 86 y los nucleótidos correspondientes que codifican para esta región de SEC ID N°: 85. La parte conservada carboxilo-terminal incluye los residuos de aminoácido de 361 a 493 de SEC ID N°: 86, y los nucleótidos correspondientes que codifican para esta región de SEC ID N°: 85. Por tanto, en resumen, estas secuencias representan dianas terapéuticas y de diagnóstico atractivas para EII.

45 La figura 1 muestra un esquema de clones de flagelina con porcentaje de similitud con respecto a flagelina B relacionada de la bacteria ruminal anaerobia, *Butyrivibrio fibrisolvens*. (A) Estructura de flagelina A de *B. fibrisolvens* que muestra regiones amino y carboxilo conservadas y el dominio central hipervariable. (B) Mapeo de las secuencias de aminoácidos predichas a partir de los clones de expresión de flagelina (1-12) en relación la secuencia de *B. fibrisolvens*. El intervalo de porcentaje (41-84%) indica la similitud en la secuencia conservada NH<sub>2</sub> entre las secuencias del clon y de *B. fibrisolvens*. (C) Diagrama de la secuencia de aminoácidos de longitud completa de flagelinas cBir-1 y Fla<sup>x</sup> de bacterias cecales de ratón que indica la similitud de los tres dominios con los dominios de *B. fibrisolvens* respectivos. (D, E) Esquemas de proteínas de flagelina recombinantes y fragmentos para cBir-1 (D) y Fla<sup>x</sup> (E) que se han expresado y purificado en *E. coli* mediante afinidad con etiqueta de 6 histidinas para columnas NiNTA (Qiagen Corp.).

#### 55 Ejemplo 5

##### CEBADO DE PÉPTIDOS DE LÍNEAS COOPERADORAS T

60 La generación de líneas cooperadoras T CD4<sup>+</sup> y la identificación de epítopos peptídicos derivados de antígenos específicos de bacterias que pueden reconocerse por células T CD4<sup>+</sup> en el contexto de moléculas de clase II de HLA, se lleva a cabo de la siguiente manera:

65 Se generan péptidos de quince meros que se solapan por 10 aminoácidos, derivados de un antígeno específico de bacterias, utilizando procedimientos convencionales. Se derivan células dendríticas (CD) de CMSP de un donante normal utilizando GM-CSF e IL-4 mediante protocolos convencionales. Se generan células T CD4<sup>+</sup> a partir del mismo donante que las CD utilizando perlas de MACS (Miltenyi Biotec, Auburn, CA) y selección negativa. Se pulsan

las CD durante la noche con combinaciones de péptidos de 15 meros, con cada péptido a una concentración final de 0,25 µg/ml. Se lavan las CD pulsadas y se siembran en placas a  $1 \times 10^4$  células/pocillo de placas con fondo en V de 96 pocillos y se añaden células T CD4<sup>+</sup> purificadas a  $1 \times 10^5$ /pocillo. Se complementan los cultivos con 60 ng/ml de IL-6 y 10 ng/ml de IL-12 y se incuban a 37°C. Vuelven a estimularse los cultivos como anteriormente de manera semanal utilizando CD generadas y pulsadas tal como anteriormente como células presentadoras de antígeno, complementadas con 5 ng/ml de IL-7 y 10 U/ml de IL-2. Tras 4 ciclos de estimulación *in vitro*, se someten a prueba líneas de células T CD4<sup>+</sup> resultantes (correspondiendo cada línea a un pocillo) para determinar la proliferación específica y producción de citocinas en respuesta a las combinaciones estimulantes de péptidos utilizando una combinación irrelevante de péptidos como control.

### Ejemplo 6

#### GENERACIÓN DE LÍNEAS DE LTC ESPECÍFICAS DE BACTERIAS UTILIZANDO CEBADO DE GENES COMPLETOS *IN VITRO*

Utilizando cebado de genes completos *in vitro* con CD infectadas por vaccinia con antígeno bacteriano (véase, por ejemplo, Yee *et al*, The Journal of Immunology, 157(9):4079-86, 1996), se derivan líneas de LTC humanos que reconocen específicamente fibroblastos autólogos transducidos con un antígeno bacteriano específico, según se determina mediante análisis ELISPOT de interferón. Específicamente, se diferencian células dendríticas (CD) a partir de cultivos de monocitos derivados de CMSP de donantes humanos normales haciéndolas crecer durante cinco días en medio RPMI que contiene el 10% de suero humano, 50 ng/ml de GM-CSF humano y 30 ng/ml de IL-4 humano. Tras el cultivo, se infectan las CD durante la noche con virus vaccinia recombinante con antígeno bacteriano a una multiplicidad de infección (M.O.I) de cinco, y se maduran durante la noche mediante la adición de 3 µg/ml de ligando CD40. Entonces se inactiva el virus mediante irradiación con UV. Se aíslan células T CD8<sup>+</sup> utilizando un sistema de perlas magnéticas y se inician cultivos de cebado utilizando técnicas de cultivo convencionales. Vuelven a estimularse los cultivos cada 7-10 días utilizando fibroblastos primarios autólogos transducidos de manera retroviral con antígenos bacterianos previamente identificados. Tras cuatro ciclos de estimulación, se identifican líneas de células T CD8<sup>+</sup> que producen específicamente interferón cuando se estimulan con fibroblastos autólogos transducidos con antígeno bacteriano. Utilizando un panel de líneas LCL-B no coincidentes con HLA transducidas con un vector que expresa un antígeno bacteriano, y midiendo la producción de interferón por las líneas de LTC en un ensayo ELISPOT, se determina la restricción por HLA de las líneas de LTC.

### Ejemplo 7

#### GENERACIÓN Y CARACTERIZACIÓN DE ANTICUERPOS MONOCLONALES ANTI-ANTÍGENO BACTERIANO

Se preparan anticuerpos monoclonales de ratón contra proteínas de antígeno bacteriano derivado de *E. coli* de la siguiente manera: Se inmunizan ratones con adyuvante completo de Freund (CFA) que contiene 50 µg de proteína bacteriana recombinante, seguido por un refuerzo intraperitoneal posterior con adyuvante incompleto de Freund (IFA) que contiene 10 µg de proteína recombinante. Tres días antes de extirpar los bazo, se inmunizan los ratones por vía intravenosa con aproximadamente 50 µg de proteína recombinante soluble. Se extirpa el bazo de un ratón con un título positivo frente al antígeno bacteriano, y se prepara una suspensión de una única célula y se utiliza para la fusión con células de mieloma SP2/O para generar hibridomas de células B. Se someten a prueba los sobrenadantes de los clones híbridos mediante ELISA para determinar la especificidad frente a proteína bacteriana recombinante, y se mapea el epítipo utilizando péptidos que abarcan la secuencia de proteína bacteriana completa. También se someten a prueba los AcM mediante citometría de flujo para determinar su capacidad para detectar proteína bacteriana sobre la superficie de células transfectadas de manera estable con el ADNc que codifica para la proteína bacteriana.

### Ejemplo 8

#### SÍNTESIS DE POLIPÉPTIDOS

Se sintetizan polipéptidos en un sintetizador de péptidos Perkin Elmer/Applied Biosystems Division 430A utilizando química de FMOC con activación por HPTU (hexafluorofosfato de O-benzotriazol-N,N,N',N'-tetrametiluronio). Se fija una secuencia de Gly-Cys-Gly al extremo amino-terminal del péptido para proporcionar un procedimiento de conjugación, unión a una superficie inmovilizada o marcaje del péptido. Se lleva a cabo la escisión de los péptidos del soporte sólido utilizando la siguiente mezcla de escisión: ácido trifluoroacético:etanoditol:tioanisol:agua:fenol (40:1:2:2:3). Tras escindir durante 2 horas, se precipitan los péptidos en metil t-butil éter frío. Entonces se disuelven los gránulos de péptido en agua que contiene el 0,1% de ácido trifluoroacético (TFA) y se liofilizan antes de la purificación mediante HPLC de fase inversa C18. Se utiliza un gradiente del 0%-60% de acetonitrilo (que contiene el 0,1% de TFA) en agua (que contiene el 0,1% de TFA) para eluir los péptidos. Tras la liofilización de las fracciones puras, se caracterizan los péptidos utilizando electrospray u otros tipos de espectrometría de masas y mediante análisis de aminoácidos.

**Ejemplo 9****ANTICUERPO REACTIVO CON FLAGELINA EN SUERO DE RATONES COLÍTICOS Y SU RELACIÓN CON LA PATOLOGÍA**

Se determinaron respuestas de anticuerpos en suero frente a flagelinas recombinantes cBir-1 y Fla<sup>x</sup> y fragmentos mediante análisis de inmunotransferencia de tipo Western. Se sometieron cien microgramos de proteína de longitud completa (FL) así como fragmentos de la región conservada amino (A) y la región carboxilo conservada (C) a separación por SDS-PAGE, se transfirieron a membrana y se sometieron a suero de ratones C3H/HeJBir no colíticos (combinación de 2) y colíticos (combinación de 5). Se realizó un segundo experimento utilizando 20 ng de proteína recombinante y suero de ratones FVB no colíticos (combinación de 5) y Mdr1a<sup>-/-</sup> colíticos (combinación de 5). Se utilizaron muestras humanas en un tercer experimento. Se ejecutaron muestras de un donante de sangre humano al azar y un paciente con enfermedad de Crohn con enfermedad grave frente a 50 ng de proteína recombinante. Los resultados de estos análisis de inmunotransferencia de tipo Western se muestran en la figura 2a. Se utilizaron antígenos de bacterias cecales (CBA) de ratón como control.

La figura 2b muestra la titulación de anticuerpo anti-flagelina en suero frente a flagelinas recombinantes cBir-1 y Fla<sup>x</sup> con anticuerpos secundarios específicos para anticuerpos IgG y IgG2a de ratón. Se utilizó suero de ratones colíticos (combinación de 5) frente a ratones C3H/HeJBir no colíticos (combinación de 2) en el panel superior y se utilizó suero de ratones MDR colíticos (combinación de 5) frente a ratones FVB no colíticos (combinación de 5) en el panel inferior.

La figura 2c muestra la correlación de la puntuación de colitis con anticuerpo anti-Fla<sup>x</sup> en suero a una dilución de 1:200 ( $r = +0,7585$ ). También se determinaron correlaciones para la edad del ratón frente a la puntuación de colitis ( $r = +0,3716$ ), y la edad de ratón frente a la DO450 de anticuerpo anti-flagelina a una dilución de 1:200 ( $r = +0,3253$ ). Las puntuaciones de colitis se basaron en la siguiente escala: sin enfermedad (0-2); enfermedad leve (3-15); enfermedad moderada (16-35); y enfermedad grave ( $\geq 36$ ).

**Ejemplo 10****LOS MACRÓFAGOS DERIVADOS DE MONOCITOS HUMANOS ESTIMULADOS CON FLAGELINA LIBERAN INTERLEUCINA 10 Y FACTOR DE NECROSIS TUMORAL**

Se cultivaron macrófagos derivados de monocitos humanos de siete donantes sanos *in vitro* durante cinco días y después se estimularon con flagelina X de longitud completa (SEC ID N°: 79) a 6, 12 o 24  $\mu\text{g/ml}$ . Se detectó liberación de citocinas IL10 y TNF- $\alpha$  en el sobrenadante de cultivo mediante ELISA. La liberación de citocinas tras la estimulación con flagelina X fue dependiente de la dosis, véase la figura 3.

**Ejemplo 11****PROTEÍNAS DE FLAGELINA EN COMBINACIÓN CON ADYUVANTE DE TOXINA DEL CÓLERA COMO VACUNA PARA EII**

Se utilizó proteína bacteriana de flagelina, Fla<sup>x</sup>, junto con un adyuvante conocido, toxina del cólera (CT), para prevenir la aparición de colitis grave en ratones que son genéticamente propensos a la enfermedad. Ratones mutantes MDR1a (ratones MDR1aKO) desarrollaron inflamación crónica grave en el colon a medida que avanzó su edad. O bien se dejaron sin tratar ratones hembra deficientes para MDR1a de 5-6 semanas de edad de Taconic Farms o bien se les administraron 8  $\mu\text{g}$  de Fla<sup>x</sup> u 8  $\mu\text{g}$  de Mtb + 10  $\mu\text{g}$  de toxina del cólera (List Biologics) en un volumen de 200  $\mu\text{l}$  de PBS mediante sonda nasogástrica. Por tanto, se trataron los ratones en el día 1 y el día 14 del experimento. Se sacrificaron los ratones 4 semanas tras la última dosis, y se fijaron sus intestinos gruesos (ciego/colon/recto) en el 10% de formalina tamponada neutra (Sigma), se cortaron en rodajas y se tiñeron con hematoxilina y eosina (H&E). Se examinaron las secciones de manera enmascarada y se puntuaron en cuanto a la gravedad de la enfermedad.

Las puntuaciones oscilaron entre unidades arbitrarias de 0 (para ningún cambio observado en ningún sitio) y 64 (la inflamación más grave y cambios epiteliales observados por todo el intestino grueso completo). Normalmente, la enfermedad leve presenta una puntuación de 1-15, la enfermedad moderada de 16-35, y la enfermedad grave de  $>35$ . La puntuación media para ratones no tratados ( $n=4$ ) fue de 23,50 con una desviación estándar de 17,37. La puntuación media para ratones a los que se les administró proteína de control más CT ( $n=5$ ) fue de 25,40 con una desviación estándar de 18,43. La puntuación media para ratones a los que se les administró Fla<sup>x</sup> más CT ( $n=5$ ) fue de 6,80 con una desviación estándar de 4,38. Las puntuaciones medias de los ratones tratados con proteína de control frente a tratados con Fla<sup>x</sup> fueron significativamente diferentes (prueba no paramétrica de Mann-Whitney,  $p=0,0159$ ). Por tanto el tratamiento de ratones MDR1aKO con Fla<sup>x</sup> más CT dio como resultado enfermedad menos grave.

A partir de lo anterior se apreciará que, aunque en la presente memoria se han descrito realizaciones específicas de

la invención para fines de ilustración, pueden realizarse diversas modificaciones sin desviarse del alcance de la invención. Por consiguiente, la invención no se limita excepto mediante las reivindicaciones adjuntas.

**Listado de secuencias**

- 5 <110> Corixa Corporation
- <120> Composiciones y procedimientos para la terapia y el diagnóstico de enfermedad inflamatoria del intestino
- 10 <140> EP 02805613  
<141> 2002-12-16
- <150> US 60/341.830  
<151> 2001-12-17
- 15 <150> US 60/396.242  
<151> 2002-07-16
- <150> US 60/426.835  
<151> 2002-11-15
- 20 <160> 87
- <170> Patente en versión 2.1
- 25 <210> 1  
<211> 1268  
<212> ADN  
<213> Desconocido
- 30 <220>  
<223> Descripción de secuencia desconocida: clon 76779 Cbir-1 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/HeJ Bir
- 35 <400> 1
 

ggaggtatta	ttatggtagt	acagcacaat	ttacatgcaa	tgaactctaa	cagaatgtta	60
ggcatcacac	agaagacagc	atctaagtct	acagaaaagt	tatcttcagg	ttacgcaatc	120
aaccgcgcag	cagacaacgc	agcaggctct	gctatttctg	agaagatgag	aaagcagatc	180
agaggactta	cacaggcttc	tacaaatgct	gaggacggca	tcagctctgt	acagacagca	240
gaaggcgctt	tgacagaagt	gcatgatatg	cttcagagaa	tgaacgagct	ggcaattcag	300
gcagcaaacg	gcacaaactc	agaagatgac	cgctcataca	ttcaggacga	aattgaccag	360
ctgacacagg	aatcgatcg	tgttgctgag	acaacaaagt	tcaatgagac	atatctcttg	420
aagggtgaca	caaagaacgt	tgacgctatg	gactatacat	atagctataa	ggcagttaca	480
acgaatactg	tagcaagagc	ttcggtttta	gcagcagaga	acacagctac	aggtatgtca	540
gttagtattt	cattgctgc	aaacagcggc	aaggttactg	cagctgactc	taacaacctt	600
gcaaaggcta	tcagagatca	gggcttcaca	atcacaaact	ctaccagaa	tggttaaggtt	660
gtttacggtc	ttgagctgaa	cggaagcgat	gcaaaggcaa	actatacagt	ttcaacagta	720
agtatggaag	ctggtacatt	caagatcctg	aattctaata	agcaggttgt	tgcatctgta	780
acaatatcta	caacagctag	ctttaaaaag	gtatctggta	tgtcacagat	cgttacggcg	840
tactctgtat	cagcagctta	tgcgacgggt	gatgtatact	ctctctatga	cgcagacgga	900
aatgcaattt	cagcaaacia	gctggataag	tactttacgg	caggcggcgc	tacagaggca	960
ggcggaatag	ctactacact	ttcagcaaac	tctgggtgtc	ctaaggttta	tgacgtactc	1020
ggaaaagagg	tttctgcagt	aagcattgca	agtactttag	taacagcagt	taaggataag	1080
acggctgcat	tgaagatgaa	cttccatgta	ggtgctgacg	gaacagataa	caacaagatt	1140
aagatcaaca	ttgaggctat	gacagctaag	agtcttggag	ttaacggctc	gaaggtgagc	1200
ggttcgagcg	gaacaaatgc	tacaaatgct	atcgagataa	tcgctggcgc	tatcaagaag	1260
gtttctac						1268
- 40 <210> 2  
<211> 336  
<212> ADN  
<213> Desconocido
- 45 <220>  
<223> Descripción de secuencia desconocida: clon 76780 Cbir-2 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/HeJ Bir

ES 2 405 405 T3

<400> 2

```

gcgcaggaaa aaagttacca gcgtggacaa ggcaaatgtg ctggattcct caaggctttg 60
gcgaaaagt gtagaagaag tgggcaaaa agtaccggga cgtggcattg gagcatatgc 120
tggtagataa ctgtgccatg cagctagtaa aagaccaag gcagtttgac gtgatcctga 180
cagaaaatat gtttggcgt attttgtccg atgaagcaag catggtgaca ggctccattg 240
ggatgcttcc ctccgccagc ttaaacgata ccaaatttgg gctgtatgaa ccaagcggcg 300
gttctgcgcc ggatattgcc gggaaagggg ttgcca 336

```

5

<210> 3  
 <211> 658  
 <212> ADN  
 <213> Desconocido

10

<220>  
 <223> Descripción de secuencia desconocida: clon 76959 Cbir-3 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

15 <400> 3

```

ggccccgcc ttcggcatgg tgggcaccct ggtgggctg atcaacatgc tgaaggccat 60
ggacatcgag accgttggcg gcaacctggg ccccgctatg gccaccgctc tggtcaccac 120
cctctatggt tgcgtgctgg cccacatgat cttcggcccc atcgccacc cagctgcgca 180
gcgggacgag gaagagacc tctgcaagct gatcatcgtg gaggggctca tgtccatcca 240
ggccggcgcc aaccccaagt tcctccggga gaagctgctc accttcgtca cccagaaaca 300
gcgtggcgag aacggcggca agaagggcaa gtaagagctg cggggccgcg tccccaccgc 360
tctgcggtat gaactgaggt gaaacacccat ggcaagcatc aagaaaaaga gctccggcgg 420
cggcggcgcc aactggatgg acacatacgg cgacatggtc accctgctgc tgtgtttctt 480
cgtcctgcct gtattccatg tccacgatcg actcggagaa gtggaagatg atcgggccag 540
agcttcaata agaacgcagt cgtcagcgac gatcagcccc ccggaccgga cggcactgaa 600
agcagcacgg gcggcatgaa tctgcctttg acccaggaca tgcaggccgc catggatc 658

```

<210> 4  
 <211> 644  
 <212> ADN  
 <213> Desconocido

<220>  
 <223> Descripción de secuencia desconocida: clon 76960 Cbir-4 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

<400> 4

```

cttgggaagt atttaaaagc gcaggcattg ctggaaaagc tgcgctatta tgcggagaaa 60
tgcgggcgca cctatatccg catggaagtg tgcctcctta gcgccgtagc gaaataccga 120
acaggcgggg aatggaagg ggaatttttc ccgatgctga gagaagcctg cggatatcat 180
tttatccgcc ttgtcagcga ggaaggggcg gcggcgcagg aactgtttgc ggccggcgga 240
aagagtcttc tggaaaagga agtaatggat aaggcatggc tgtccaggct catggaggaa 300
acagggaagg tggcgggtgc ctacctggcg tatttaaaag gccggcttgc cgaagcggcg 360
gatttctgtg aggcggcatt atccatcctg cgctgcagg cggaaggaaa gagcgcatca 420
atgaagcttg gagttttttt tgttcggcta ttttttggcg tcgtaagct ttacgttttt 480
gaaagtcagc acagagaaag gagaggtagg cagcggagta taagacatgc ggaagcaaat 540
gaaccaaatg gaggtagaaa catgacgtta tttcaatgaa cagttgtttg aagcaggtgt 600
tctttcggac atcagacaag aagatggaac cctaagatgg ctcc 644

```

<210> 5  
 <211> 685  
 <212> ADN  
 <213> Desconocido

<220>  
 <223> Descripción de secuencia desconocida: clon 76961 Cbir-5 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

40 <400> 5

ES 2 405 405 T3

```

aatgaaactt atcaaaaaa atctctccat cacacaaatg cttgttgaaa catcatctca 60
tcttgatagc aacctaaaa atattgcaa aatttcacag gataacaccg agctaggcga 120
aaggagtgtg aatatcattg agcaaaacat caccctttca aatgcaacaa aagaatcttt 180
agaagatgtg cttaatacga tgcagcaaac tcaagcactc ataagctcta tcaacgaaga 240
aatcacaata gacgcacaaa aagaagatga aaatatgcaa aagattctct ctcttgctaa 300
tgaggcaaaa aatattcaaa gtgtacttgt aactattacc gacatagcag accaaacaaa 360
cctcttagca cttaatgcag ctattgaagc agcgagagct ggggagcacg gacgaggctt 420
tgctgtgtgt gctgatgaag tgagaaaact agcggagcgc acacagcatt ccattaccgg 480
agacaggtag cattatcaa tctgtcttgc agtctattga tgaagtatca agtgatatgg 540
gaaaccagtg ccaaatcaat gaataatctt tcaagcaggg tgaagtgatg ttggcaatat 600
acaatctctt gccactcgt gcaagaaaac aatgcaaaat cattgcaaag gctagaggga 660
gcccgaatgg cgaatgaaaa tacca 685

```

<210> 6  
 <211> 929  
 5 <212> ADN  
 <213> Desconocido

<220>  
 <223> Descripción de secuencia desconocida: clon 76781 Cbir-6 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

<400> 6

```

atcatttcta cgaaccaggg tgtcatcaca gacaaagaag caagaaagct cggcgtaggc 60
ggagaagtac tggcatttgt gtggtaggca ctggcactga acacttagcg aagccaagcg 120
ttccgcgcag gctgaggtta gtgtgaaggc cgttctcaga gcgaagcgaa gagaacttcc 180
ttgctaggga acacttagcg aagccaagcg ttccgcgtag gctgaggtta gtgtgaaggc 240
cgttctcaga gcgaagcgaa gagaacttcc ttgctaggga acacttagcg aagccaagcg 300
ttccgcgcag gctgaggttg gtgtgaaggc cgttctcaga gcgaagcgaa gagaacttcc 360
ttagaggaaa caccatcgtc actagactcg aaagaaactc gccaaagtgt cacagcaact 420
gaaaacagag caggcgaag gttctgctcc gagaatttaa gtaaggagg atatggcaat 480
gtcacgtatc ggaagactgc caatcgcgat tccggcagga gtaactgtgg aaatcgcaga 540
gaataatgta gtgaccgtaa aaggtccaaa gggaaactct tctcgggagc ttctgttga 600
aatggaaatt aagaaaagac gcgagacaat cgtcgttaca agaccgaatg atttgaagaa 660
gatgaaatcc cttcacggcc ttaccagaac actgattaac aacatggtta tcggcgttac 720
agaaggatat aagaaagttc ttgaagtaaa cgggtgttgt tatagagcag caaaatcagg 780
aaacaaatta acacttagcc ttggatattc ccatccggta gagatgatcg acccggaagg 840
cgttgagacg gttctcgagg gacagaacaa gattaccggt cagggtatcg acaaggaaaa 900
ggttgacag tatgcagccg agatcagag 929

```

<210> 7  
 <211> 652  
 <212> ADN  
 <213> Desconocido

<220>  
 <223> Descripción de secuencia desconocida: clon 76962 Cbir-8 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

<400> 7

```

gcatcaaaaag catcgtaat tgcgcctcct gctggaacca tagggaaaac tttatcatct 60
tcttcaatca cgcattcaat tactactggg ttctttaatg caatggcttt ctcaatcgca 120
ggcgcaacct cctcacgctt ggtaaccggg attgcctcac agcctaaacc ctctgaaacc 180
tttacaaaat cgacctatc cgtagaatg gtctgtgaat accgtttacc ataaaataaa 240
gtctgccact ggcgcacct acccaaaaaca tgattattta tcacaatctg gataattgga 300
atattatagc ggcttgcaat tgccaactca tcaaatcca tccgaaaaca cccgtcacct 360
gcaatattca cacatattt atctggtctt ccaacctttg caccgatata tgetccaaga 420
ccatatccca ttgtaccaag acctcccgaa gttaaagaaag tacggggttc tgtataccta 480
taaaactgtg ctgccacat ttgatgctgg cctaccgct caatttttg ttgctatgt 540
agctcagggg agaagcactt cttgtaagg gaagaagtcg cggggttcaa aatcccgcca 600
tcggctttta agtaaaaaac cccggggagg ggcttatttg cattaatccc cg 652

```

<210> 8  
 <211> 497  
 <212> ADN  
 <213> Desconocido

ES 2 405 405 T3

<220>

<223> Descripción de secuencia desconocida: clon 76782 Cbir-9 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

5

<400> 8

```

cttcgctcgg atggcttcat cctcgaccgc gtcacaaatg ggccctgggt gtctactact 60
gctggttctg ctacggacgg tcacctcttg aatacgtacc cgacgaatat ctcccctcag 120
gataaccgtt ccccggggtt cggttttgcc gttcgctgtg tggtagcgga ggggtggagg 180
ctaaatctgc ttctacgcg tcgctgggcg tttgcgtggc ggtacggcat ttttcttagc 240
agccccgcgc gtagccggac tagacactcc cgttgcgggt ttccctgcgc cccccctgt 300
tgttcgagcg gttttttag ctcctcttgc tgccgctcta ctattgcttc tcgtagcctg 360
acttctagtt gtttttctac ttgatgccct agcgttttcg gcgatttgct tctcgattcg 420
tgagtaaatg ctacgagcgc cctcgggtgc acccattgcc tcataaatcg ccaaaatttg 480
ccgcaaagta tttgtac                                     497
    
```

10 <210> 9

<211> 327

<212> ADN

<213> Desconocido

15 <220>

<223> Descripción de secuencia desconocida: clon 76963 Cbir-11 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

<400> 9

20

```

gtgtgtggtc cacggcgcgg ccttcaccgg cggtagcgc acggaccagg tgctggcgga 60
cttcaccgcc ccggaggacg gtctcttcca catcctctgc ctccacggcg acgtcttcag 120
ccaggacagc gtctacggcc ccataccccg gccccagatc gcccgacgc gcgcggtata 180
cctggccctg ggccacgtcc accagtgcag cggcatccag cggcaggggg acaccccctg 240
ggcctacccc ggctgtcccg agggccgggg gttcgacgag ctgggggaca aggggtgtgct 300
ggcgggggac gtggatcggg gcggggc                                     327
    
```

<210> 10

<211> 621

25 <212> ADN

<213> Desconocido

<220>

<223> Descripción de secuencia desconocida: clon 76964 Cbir-12 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

30

<400> 10

```

attgagatac cattggccga gtgattgagg cttatagga taaacaaaaa cgtgtctgag 60
atgttactta taaagtgtct ttgtgcctgc aggtgtcatt tctcgaacct ttattgttaa 120
atataataaa taaggaaaaa tttctttaat aatacgtcat ataccgcaa attttagtaa 180
ctttgcactc gggaaaaaac gtgttttta gaaatcaaaa tttatcaact cccaataaat 240
aacaaggtca tttactacaa tgagcacaat tgaagaaatc aagaaagccc gtgtagccga 300
catcaagaag aatcttgaaa gctacggcat cgacggcaca actgaaatcg tgtacaacc 360
tacctatgag cagctctttg ctgaagagac actccccttg ctcgaaggat atgaaaagg 420
tgttgccact gagcttgacg ctgtcaatgt tatgaccggc gtatataacc gccgttcacc 480
caaggacaaa ttcactgtac tcgacgaaaa ctcaaaggac accgtatggt gggacaaccg 540
aagaatacaa gaacgacaac aagcccgtt ccgaaagagg catggaaggc ttgcaaggaa 600
cttgcagtga aaggaaacct a                                     621
    
```

35

<210> 11

<211> 625

<212> ADN

<213> Desconocido

40

<220>

<223> Descripción de secuencia desconocida: clon 77529 Cbir-13 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

ES 2 405 405 T3

<400> 11

```

acggtacggg caagcgggat cgcgacttta ggatctgcaa gtcccaaagc ctttgccgct 60
ttcagcgcgt cctttcctac agactgtgcg gctgtagcag gacggctacg catcatctca 120
atctcgttat gaatctgagc gcgctgctcg ccggcgcgta cggtcagttg ggcaagtctg 180
cgcgcgactt cgtcatttgc ctgcattatc ggcctccttt gtctgaaccg aaggaaatag 240
agcagataaa aaaagaaatg agggaagcgg gcattaaata aaagaatgtg accgattcgg 300
acaccgaaaa aacaagaaag agaggattta aaaatggcag gatttgacct taatagtttg 360
ctgaatgaa agagcaaagg ggcagcagga cagaagcagg aaacggcggg agcagggcag 420
gggccggcag aggggagga aagcagtttt gaggttgtaa tgcttgatgt agaggactta 480
atgccaagca aggataattt ctacacaacc gaggaataa acgagttggc ggacgctatc 540
gagttgtcgg gcggtatcga gcagaattta attgtaaagc cgggaagcac acggaaagta 600
tgaggttatc gcaggacacc gcgga                                     625
    
```

5 <210> 12  
 <211> 542  
 <212> ADN  
 <213> Desconocido

10 <220>  
 <223> Descripción de secuencia desconocida: clon 76965 Cbir-14 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

<400> 12

```

gatcgtgcc tgtttgactc ctgatattgg tcttcaacgt gagcttctct ttgtgcctgc 60
acgaggagga ctaggtgaac ggatggagat ccaggcaaat aacgtgtgtg ctcggtatggc 120
tcgacggact ggtgggaaaa gccgttcgct ctatgtgcc gagcaagtaa gcgagagtac 180
gtatcgtcca ttgttaaaag aacctgctgt tcaagaggtc gtcaatttga tcgggtcaaag 240
taatgctgtg atccacagca tcggaacagc aatgcatatg gcacatcgac gttcgtatggc 300
gccagaagtg atcgcaatgc taaataagaa aaaagcggtc ggtgaagcat ttggttactt 360
ttttgatgaa aaagggcaga tcgtgtatcg gatctcacgt atcgggatcc agttagaaga 420
cctccttcta tggaaatgtg aattgccgt cgtggtggc acttcaaaag ctaaagcgat 480
ccgtctctta catgaaacac ccttctaaac aaacattgtc tgatcacaga ccaaaggaac 540
ca                                     542
    
```

20 <210> 13  
 <211> 585  
 <212> ADN  
 <213> Desconocido

25 <220>  
 <223> Descripción de secuencia desconocida: clon 76966 Cbir-15 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

<400> 13

```

gggctgcgga ttattaaacc ggggagatat aagagtggat gccgaaaaat ggaatcttta 60
atctccggtc atagatgaaa aaaattgaa atagaaatag aagatgtaa gggtgacgaa 120
aggattcgtc gagaactatg aactgtcat agtcagccaa cacaaggagg taattttatg 180
gtagtacagc acaatcttac agcaatgaac tcaaacagaa tgtaggaat cacaacagga 240
agtttagcaa aatcagcaga aaaactgtct tcaggatata aggtaaaccg cgcagcagat 300
gatgcagcag gacttgcaat ttcagaaaag atgagaaagc agattagagg tcttacacag 360
gcttcaacca acgcagaaga tggatttagc gcagtacaga cggctgaggg tgcattgact 420
gaagttcatg atatgttaca acgtatgaa gagttggcag taaaagctgc aaacgggtaa 480
tgtctcttct tgacagacag accattccaa gatgaagtga cacagcttct cacagaagtt 540
30 gcccgtgtag cagaaacttc caaatttcaac cgaaatttat tcttgg                                     585
    
```

35 <210> 14  
 <211> 916  
 <212> ADN  
 <213> Desconocido

40 <220>  
 <223> Descripción de secuencia desconocida: clon 76967 Cbir-16 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

ES 2 405 405 T3

<400> 14

```

gtagaagatg gaacgttta gaccattcac atcacattaa aatatggtgc agataagaac 60
gaaaaagtaa ttcccgccct taagagaatc tccaaaccgg gtctccgctg atacgcaaac 120
agcgaagaga tgccgaaggt actgggcgga ctccggaattg caatcgatc tacaataaa 180
ggtgttgta ccgacaaaga agcaagaaag ctccgctgag cgggagaagt tctttgcttt 240
gtgtggtgac cgaagcact tggatgaatc aagcgaagc gagagatgaa cttagtgcga 300
ggtcgttctt tgagcttagc gatgattga atacttggcg aggtattatc gagcctagta 360
tgaaagcaga tattccgaac gaagagagga atatcttcga cagagcttag cgaaaagaac 420
agataaactg aaaacagagc agactcaagc ctctgctccg aaaatttaag ttaggaggac 480
aaaggtatgt cacgtatagg aagactgccg atcgcagtc ctgcagggtg aactgtagag 540
attgctgagc ataatgtagt gaccgtaaaa ggtccgaagg gaactctcgt aagagaactc 600
ccggttgaaa tggaaatcaa gcaggaagcg gaagaaattg ttgtcaccag gccaaacgac 660
ttaaagagga tgaatccct tcacggcctg acacgtacac tgatcaacaa catggtaac 720
ggtgtcagcc agggttatga aaagttctg gaagtaaacc gtgttggtta cagagcagcg 780
aagtcggca acaagctgac cctcagcctt gगतattcac atcctgttga gatggttgat 840
ccggaaggga tcgagcagt tctggaagcg cagaacaaga tcacggtaaa agggatcgac 900
aaagaaaaag ttggcc 916

```

- 5 <210> 15
- <211> 625
- <212> ADN
- <213> Desconocido

- 10 <220>
- <223> Descripción de secuencia desconocida: clon 76968 Cbir-18 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

<400> 15

```

aatgaatgca aacagaaact tggggatgac cacaaccgca caggcaaaat ctacggagaa 60
gctttcttct ggttaccgga tcaaccgcgc ggcggacgat gcggcgggccc tttcgatttc 120
cgagaagatg cgcagccaga tcccgccgct gaagcaggct tccaccaatg cgcaggacgg 180
catttccctg attcagactg cagaggggtg gttaaatgag cagcattcga ttttacagag 240
aatgcgcgag ctgtccggtc aggcggcaaa cgggtgagg acggacgacg accgtgagc 300
agtcaacaac gagatcagcc agctccagtc cgagctgaca aggatttccg agacgaccga 360
gttcaacacg atgaagctgc tcgacggaag cctttccggg acagccggat cgtccaccgg 420
ctcaggcccg aagttcggcg tggtagatgc aaccttagac ggcgcgcttg tgacgtcgaa 480
tgttgcaggt gtcaaggtag ctacagcttg ctacgacaag tacaagaagc ggacagggag 540
actgccatct gggacgcaac ccggaagac cttgacattg aatctttctc gtgcccgaat 600
tcaaaaaagc ttttcgaag aagta 625

```

- 20 <210> 16
- <211> 603
- <212> ADN
- <213> Desconocido

- 25 <220>
- <223> Descripción de secuencia desconocida: clon 77530 Cbir-19 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

<400> 16

```

agggcatctg aaaatcggac acaatgtgaa aatcggatat ttcgccaga atcaggctca 60
gcttctcgac ggcgaactta ccgtttttga caccatcgac agggtggtg tcggagatgt 120
ccgcacaaaa atacgcgaca tcctcgccgc attcatgttt ggcggagagg cgtcggacaa 180
aaaggtgaag gtgctctcg gggcgaaaa gactcgcctt gcgatgataa aacttctgct 240
tgagccggtg aacctgctta ttttgacga accgaccaac cacctcgaca tgaagacaaa 300
ggatatactc aagcaagcga taaaggactt caacggcacg gtgatagttg taagccacga 360
ccgcgaattt cttgacggac ttgttgaaaa ggtatagag ttccggtggt ggggctgtgc 420
gtgaaaacct tggcgaatc tacgatttcc ttgaacgcaa gcgccttgct tcattgacgg 480
agttggagcc gaaatgcacc ccgggcaaaa gatgacaaaa ctctccccg gcgaagcgag 540
aacactcagc ccgccacgg actcgaacc tttaaactac cgccgaggcc cgcgaacccc 600
gac 603

```

- 30 <210> 17
- <211> 548
- <212> ADN

ES 2 405 405 T3

<213> Desconocido

<220>

5 <223> Descripción de secuencia desconocida: clon 76969 Cbir-20 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/HeJ Bir

<400> 17

	gtctatatga	ttattcaaca	taatatagcg	gcgattaact	cttatcgtaa	cttaggcgta	60
	aaccagagcg	gactgaacaa	aaacttagag	aaactgtcat	ctggttacia	aatcaaccgt	120
	gcaggcgatg	atgcagcagg	tctggctatt	tccgagagca	tgcgttctca	gattaacggc	180
	ttaaaccagg	gaggactgaa	caaaaactta	gagaaactgt	catctgggta	caaaatcaac	240
	cgtgcagcg	atgatgcagc	aggtctggct	atttccgaga	gcatgcgttc	tcagattaac	300
	ggcttaaac	aggcagtaaa	caacgcaaa	gatgcatcg	gtttgattca	gacggcagaa	360
	ggtgctctga	ccagggattc	ggcgtatccg	tttccagcgg	tatgctgcag	gattccatcg	420
	ctgatacgaa	acggattgaa	tacacaagag	ttgaataaat	gaatagagca	gcgttgacgg	480
	gttgtgccc	g	g	g	g	g	540
10	aaataaac						548

<210> 18

<211> 535

<212> ADN

15 <213> Desconocido

<220>

20 <223> Descripción de secuencia desconocida: clon 76970 Cbir-23 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

<400> 18

	agctacgttg	agtaccagac	tgcaaatcgt	cactatgcac	acgttgactg	cccgggcccac	60
	gccgactatg	tgaagaacat	ggtaactggg	gcagctcaga	tggacgggtg	aatccttgtt	120
	tgtgctgcaa	ctgacgggcc	tatgcctcag	acacgtgagc	acatcctcct	cgcccgtcag	180
	gtgaacgttc	caagaatcgt	tgttttcatg	aacaagggtg	accttggtga	cgatcctgag	240
	atgcttgacc	tcgtagagat	ggagctccgt	gacctccttt	cattctacia	cttcgacggg	300
	gacaatgctc	cggtaatccg	tggctctgca	cttgggtgcac	tcaatgggta	ccctcagtg	360
	gaggataagg	ttatggaact	catggcagct	ggtgacgagt	atattccgct	tcctccacgt	420
	gacaacgaga	agccattcct	tatgccaatc	gaggacatct	tctctatcac	agggcgtggg	480
	actgtagcaa	ctggacgtat	cgagaccggg	gatcattcac	cgtaggtgat	gaagg	535

25 <210> 19

<211> 505

<212> ADN

<213> Desconocido

30 <220>

<223> Descripción de secuencia desconocida: clon 76971 Cbir-24 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

<400> 19

35	aggaaatcca	atggtgccc	gactgggtat	ctcagtatga	gaattacgcc	atgtatcgga	60
	aatacgggct	gggagcggta	attctgaaag	agggggagcc	ggtgtccggc	gcgtcctcct	120
	ataccgggta	tatcggcggc	attgagatcg	aaattgatac	cagagaggac	tgccgcagaa	180
	agggcctggc	ctatatatgc	gccgccagac	tgatcttgga	atgtctggac	cggggctggt	240
	atccaagctg	ggacgcgcaa	aatctgtggt	cggtggcgtc	ggccgggaag	ctgggatatc	300
	atcttgaccg	ggaatatacg	gcataatgc	gggtcaggta	ggggaaagg	aacgtaaaat	360
	caagtaccgc	gagatttgag	agaacataaa	aaataacagg	aggttccgat	agatgaaagt	420
	tttaatgctt	gaacggcagc	ccccgtgcca	acgggaatac	atacgcgctt	ctgccgggta	480
	aatggaatac	aaatcttga	acagc				505

<210> 20

<211> 532

40 <212> ADN

<213> Desconocido

<220>

ES 2 405 405 T3

<223> Descripción de secuencia desconocida: clon 77073 Cbir-26 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

<400> 20

5

gtaaggtaat	tctgttaaaa	ttgactgggg	cgtgtttgga	atgttttttg	gcagatacca	60
aacgcgccct	aagcaagaca	ggggaatgga	atatgagcgt	acaggataaa	atagaacatg	120
tattaaaatg	catacacctg	ctgttttcca	aaagccagcc	ttacggggac	agcaatacaa	180
agattattgt	ggataaaaa	gccgttttg	aactgctgga	acagttaaac	cttgcagttt	240
atgaggctat	ggaccattat	gaggtacca	cccgtaagca	tgagattgca	gagcggcgct	300

gcgagaaacg	gggcgaggaa	atcatccaga	aggccagcaa	gcacgcggac	gacatctatg	360
ccgcttccat	tatgtatacc	gacgatgcca	tcaaccggat	ttgctatatt	atggatgacg	420
cccatcaggc	tgtccaaaat	attttccgta	agatgaacgt	agagatggaa	aaaccctcgc	480
tgccgaattc	ggccgaagga	gtggggttgg	taatcgacgc	aaagccgacc	ca	532

10 <210> 21  
 <211> 467  
 <212> ADN  
 <213> Desconocido

15 <220>  
 <223> Descripción de secuencia desconocida: clon 76972 Cbir-27 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

<400> 21

20

cttacacaag	agcctgccga	aatcttaaat	cacacctact	aatatacaat	gcgtgaggat	60
attgtgatgg	caatggagga	actggagctt	acgccgcagc	aggcaaaggc	actcttgaaa	120
tctccctgtc	cgcttgatga	tgtgtataag	gaatttaagg	acagagaggt	cgagcatatg	180
gatacgattc	gtgattccat	tgaaacgaga	gcagatcagg	ttatcaagcg	ggaaaacgca	240
agggaaagca	ggtgatccta	tgccgtatat	cccaccggag	gtgattgaac	aggcaaggca	300
aattgatttg	cttcatata	tgaaagcctt	tgagccgaat	gagctggtea	ggatttccgg	360
caacaactac	accaccgcga	ctcacgacag	cttaaagatt	tcaaaccgga	aatggatgtg	420
gtggtcgcag	cgaatcggcg	ggtataacgc	ccttgattat	ctcataa		467

<210> 22  
 <211> 701  
 25 <212> ADN  
 <213> Desconocido

<220>  
 <223> Descripción de secuencia desconocida: clon 76973 Cbir-30 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

<400> 22

30

cttttgac	ttttttgat	gcaacataaa	atttttatcc	aggagaatgc	ttaatgaaaa	60
acaaaaaca	aaccataaaa	tgctcagaat	gtaaccattg	cagcggtctg	cgctcgtccc	120
gcaacacaag	cacgagcttc	acctgctccc	accctgacca	ggggtatatac	caggactatt	180
tccgtgaaa	aaggatgaac	aagatgcctg	gtttccttgg	atacggggca	agatattcgg	240
agcccgctcc	tatcaagaca	gctcccgcct	ggtgccctga	aaaaagggcg	gcaaaacaaa	300
aacgaaacgg	gcaatgaccg	ctttagaaac	aaccggggag	aggcagttta	tgctgtttct	360
cctcttttct	attccactct	atctgcataa	aaaactcaca	tacagaatat	gaaattgaat	420
tttgacata	attctgttat	aataatacac	aaataggcga	ttgcgaaaca	agttcgcaat	480
attaattcca	acagggaaaa	attcggatcat	gtcggatttt	aaggaggttg	atatgagtaa	540
tcggatttta	agacatgaaa	ataaagcaga	tagggacatc	tgccctgcgaa	atggcatggt	600
ggtatgtgaa	ggtatgcaat	cagagggttt	agactgcgga	tatatattgat	gcttgggcag	660
tttggtcagc	cgcttattac	cagtataag	ctgggtaatc	c		701

35 <210> 23  
 <211> 68  
 <212> ADN  
 <213> Desconocido

40 <220>  
 <223> Descripción de secuencia desconocida: clon 76979 Cbir-32 de biblioteca de expresión de cizalladura al azar

ES 2 405 405 T3

de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

<400> 23

5 catggggggt ttccaacacg ccgttttcca ccatgtcctt ggcgcccage aacttttct 60  
cggcgggc 68

<210> 24

<211> 572

<212> ADN

10 <213> Desconocido

<220>

<223> Descripción de secuencia desconocida: clon 77074 Cbir-36 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

15

<400> 24

tgattatcga atataatcat aaatagctga ttttagtttt tatcaataag acagcccata 60  
tgggcagaaa attaaggagg acgttttaaa atggctaaag ctaaatttga gagaaacaaa 120  
ccacattgca atattggaac cattggtcac gttgaccatg gtaaaacaac ttaacagca 180  
gctatcacia aagttttatc tgagaggggt gcaggaaaac aagctactga ctttgaaaac 240  
attgacaagg ctccagagga aagggaaagg ggtatcacia tctctaccgc acacggtgag 300  
tacgagacag acaacaggca ttatgcacac gttgactgcc caggccatgc tgactatgta 360  
aagaacatga tcaactggcg tgcaacagat gacggcgcta tccttggtag tgggctgcta 420  
cagaccgggc ggttatgggc ttcaagacaa aaaagaacct tattcccttt cttgtccccg 480  
tccagggtaa ggggccggtt tccctttaat tattccgggtt ggaattttcc attgaaaacc 540  
caaatggtg gaaccattgg gggtttggga cc 572

20 <210> 25

<211> 442

<212> ADN

<213> Desconocido

25 <220>

<223> Descripción de secuencia desconocida: clon 77531 Cbir-37 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

30 <400> 25

cttttatac gacttccatg ggcgaacgtg gtgacgtcac cgtggaggaa gtacgggtccc 60  
tcctggagaa ccacttctcc ggtggactgt cgcgcaactg caccggtgag gacatccagc 120  
gggcggtgga ggaatattac aaggtaaacc atagcgatct cgtcggcccc gctcgcagcc 180  
gcaccatcgt gcatccccgt catatcgccg tgtacctctg ccgtcagatg cttgatatgc 240  
cccagggcga tategggaag aagttcaacc gggaccactc gaccgttatc cactccactc 300  
gaaccggtga ggcgatgctc gaggacaaca tggaggttca gagcgacgtg gagaggctca 360  
tgaagatcat tcgggaatcc taagtgaaaa actgggggat tgtagggttt atcattataa 420  
ttatggcggg taccgtttat gg 442

<210> 26

<211> 566

35 <212> ADN

<213> Desconocido

<220>

<223> Descripción de secuencia desconocida: clon 76975 Cbir-39 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

40

<400> 26

gagataatac caagaaacgg cagggaaacg gaaaacaggg gagctatatg agaatacagc 60  
ataatatctc ggcaattaat tcccaagaa atattatgac aacggatcga gcattatcta 120  
agaatctgga gaagctaagt tccgataca ggattaaccg cgcgggggac gacgcggcg 180  
gccttgccat atcggaggct atgcggaacc agattaccgc gatgaatcag gggcatgag 240

ES 2 405 405 T3

```

aatgtgcagg acggaatttc cctgggtgcag acggtagaag gcgccctgac agaagtccat 300
accatgctca accgcatgaa gggcatggcg gtgcaggccg ccaacgggac atacacggaa 360
tcagaacggg ccatgttaaa ctcagagatg gaagagctaa aggcagagat tacacgtatt 420
ggtgaatcca caacgttcag cggggtaccg ctctttacga acggtgggat taagaggaac 480
gtcacgctga cctcctatta cggctgcacg ctggatctat ccagagggcga agtccacggt 540
aattattccg gcagtgtggg cagagc 566

```

<210> 27  
 <211> 577  
 5 <212> ADN  
 <213> Desconocido

<220>  
 10 <223> Descripción de secuencia desconocida: clon 77075 Cbir-40 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

<400> 27

```

tattcaaaat gtgcagagaa gggaaagtga aaaatgaaga taaacacaaa tattgcagcg 60
atcaaagcga gtggacactt gaaccggaca gaggataaga tcaactacaag cctggagaga 120
ctttcttccg gataccggat aaataaggcg gcagatgatg cggcgggaat ggcgatttcc 180
cagaaaatgc atgctcagat tcgctgatta cagcgtgctt ccagaaatgg tgcggatggt 240
atttccttta tccagaccgc agaggcgcc cttatcgaag tagaaaatat gctgcagaga 300
tgccgtgagc tttctgtgca ggcggctaac caggggtgga ccatgttggga agataaagag 360
gcgatccaga aagagatcga ttctctgatg gaagagattg accgtctttc aactgatacg 420
gaatttaaca cgaaaagtat tctggatggt tctgctgtc gtcagacatc ttctaataat 480
attggagtga aggttgtttc catgacagac tcggtgggaa atgacaaaat attggaatgg 540
ctttggacca ggtggccacg aagaccattt tttccat 577

```

15 <210> 28  
 <211> 361  
 <212> ADN  
 <213> Desconocido

20 <220>  
 <223> Descripción de secuencia desconocida: clon 76976 Cbir-41 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

25 <400> 28

```

ggaggggctg aaggagggtg cccaggggat cgccgcgcg gagcagcggc ggacgggtga 60
taccgtggcc cagggggaca ccctgtgggg ggtggccggg cggtagcggc tgaccataga 120
ggcgtgctg cgggccaatc cggccatcaa gaacccaac ctgatccggg tgggacagca 180
ggtggtggtg ccggtatgac agtgcggctc ttgacgggtg acggccggca gtttgagctg 240
ccggtgacc tccggtggcg tctcctgct accggcggcg tgccctgcga cgagatggag 300
gcggtgtgcc tctacgatgg gaagctgggg gcaatcctgc ccctgtgcca ccggttcgcc 360
c 361

```

30 <210> 29  
 <211> 596  
 <212> ADN  
 <213> Desconocido

35 <220>  
 <223> Descripción de secuencia desconocida: clon 76977 Cbir-44 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

<400> 29

```

gcagcagaaa gactttcgtc cggttacaga gtcaaccgtg cagcagatga tgcggcagca 60
atggcgattt ctgagaaaaa acgggcacag atacgagggc ttgcccgcgc ttcaaaaaat 120

```

ES 2 405 405 T3

```
gcacaggatg ggattagttt tgtacagact ggagatggag ccatgagcca gattggggca 180
atgctccacc gaatgcggga gctcacggtc caggcgttga atgacggggt ctacgaacca 240
gctgaccggg cggcgctgca gatggagttt gatcagctgc aggggtgagat tgaccgggtg 300
aacgaccaga cagaatttaa taaaaaacct gtatttgaac attatacaga caatttttca 360
ttgcttgagg gaaaccgggt ttggagccag gatcagatcc ataccatcga tagcagcaat 420
tcctctctca cggtgaaata tattcggtg gagccagacg ggacagaggt agaaaaagaa 480
aagacactta cgattcctga ggggacatat accacccggg aattgatgga tgaaatggat 540
aatgtggtgt cagcgctggg agatgagggc gatggactgt atctggaata tagtgg 596
```

<210> 30

<211> 389

5 <212> ADN

<213> Desconocido

<220>

10 <223> Descripción de secuencia desconocida: clon 77532 Cbir-45 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/HeJ Bir

<400> 30

```
aaccgaagcc ggtctggaag tctatggcca cactctatgc tggcatgaac agcagaccgt 60
gaagtggctg aactctctaa tcaaggataa ggagcttgag gtggatccta acgagaagggt 120
ggaaaaagag gattatgtta tggattattc tactgtttca tttacaatt tctgggctcc 180
cgatgaagtt aagccaaca ttacgataaa tgagggttgc tttgaacttg taaaccgac 240
tgctactgac aattggaaga tccaatatca tgttgccgac ggacttccga ttgaaaaagg 300
aaaatcgtac aggcttaaag tgatggcccg tggtagccgt gaaggtacga tggaaagtaa 360
agtcggtgat tggggaggcg gatcaaatg 389
```

15

<210> 31

<211> 592

<212> ADN

<213> Desconocido

20

<220>

<223> Descripción de secuencia desconocida: clon 77533 Cbir-46 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/HeJ Bir

25 <400> 31

```
agaaacatgg gtcttcttca tggaaagact gctagacca cactctcttg tgccactgca 60
tgatgactga ttagaatgag agccatgggt agcagcctcg tgccggaaaa caacacttaa 120
cataccttag atttacgcga tcatgaaaac aaaacgactt ttactcgggt acgaggcgtt 180
tgcattaggt gcaatcaatg ccggtctgtc aggcgcataat gcatatcccg gcacaccgtc 240
cactgaaata atggagtatg tgacagacaaa tcctgtggca aaggagcgag gtatacattc 300
ccatttgtca tcaaacgaga agacagcaat ggaagaagct ctccggcatgt cattctgcgg 360
taagcgtaca ttcacgtcaa tgaagcatgt ggtctgaaat gtagccgccc atcctttcgt 420
caattctgca atgacgggag cgaacggcgg tcttcttgtc gtggcggcag atgatcccgg 480
aatgcattcg tcgcaaatg agcaggattc gcgtttttat gctgattttg cgatgattcc 540
cgcacttgag ccttcaagtc agcaaggagg cttatgatat ggcgcgtgcg gg 592
```

<210> 32

30 <211> 483

<212> ADN

<213> Desconocido

<220>

35 <223> Descripción de secuencia desconocida: clon 77534 Cbir-49 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/HeJ Bir

<220>

<221> base modificada

40 <222> (362)

<223> n = g, a, c or t

<400> 32

ES 2 405 405 T3

```

ggggcttgcg acgaccacgt cgaagtcgaa ccagttctcc ttctggatct tctccacca 60
ttctgtattg gtatcattga gagattgaat cgtaattcg gtcatgcggt gaagaatc 120
atgtgcttca tttaaagcac catctccgt ctgtaccag gagatgccat cctcgacatt 180
cttgcttccc tggtttagcc ctcgaatcat atagcgcat ttttcagaaa tggcaagtcc 240
tgcggtatca tcagcagctt tgttgatacg atagccagag gaaagcttct cgacagactt 300
ttgtttctta ttggtgtga tattcagttg atttgccgta ttaaggcag atagattagt 360
gntgacaatt aacatattta gccctcttt ctaaaataat catgttattg caaatccatc 420
gccaatatta taaattttg agtttccca ttttgacat aaaagcgcga tggggtcgct 480
ccc 483

```

<210> 33  
 <211> 214  
 5 <212> ADN  
 <213> Desconocido

<220>  
 <223> Descripción de secuencia desconocida: clon 77535 Cbir-50 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

<400> 33

```

gcttcaagaa aatttataa attttttaag aaaaaataga tttgccaca atgccacg 60
agtatgtggt attatggtat tgtcgaaaga caaggaaga gaggtattta gatactcctt 120
ttataagcac tcaaactggc gatagaaaa gcctctcaa aggggtgctt ttctgttga 180
gagaaacagg tgggaagggt gtgtcgcggc tcat 214

```

15 <210> 34  
 <211> 375  
 <212> ADN  
 <213> Desconocido

<220>  
 <223> Descripción de secuencia desconocida: clon 77076 Cbir-61 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

25 <400> 34

```

tatacgaga gcttgtccag gcttttacgg tttgccgaa accgcatgga agcggtgagg 60
ctataccaga tgatactca ggttttgtgg ctgattctgc tgtttgcagg gatcagcctg 120
ctctttggcc gcgtggcagg aatcgtgacc ggaagtgttc tggcggtttc tttctggatg 180
atggagacca ttcttgaat ttggcccga aacttctata tgctgcattt tactatagcg 240
cttatgttct taggatcgt cagataccgg atcaaaaaag ggggatggcc ctcaaacgat 300
tttgccggc tgtgcctggc tgccatcggg ttttatgtgg gcgttctctg catttgggat 360
attctggggg gcata 375

```

30 <210> 35  
 <211> 481  
 <212> ADN  
 <213> Desconocido

<220>  
 <223> Descripción de secuencia desconocida: clon 77536 Cbir-62 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

<400> 35

```

attgcaatcg tgatgaaccg tacgtgtggt aagctcacgt ttcgaagatg ttgtgtatta 60
ctctattctg ccacttaaga ttctcaccga tacaagcacg ggcacaaatc tctttgaagt 120
ctatacaaaa agccttaag aaggattac aggcactcaa gaatcgttgc aaaatgttct 180
taaagaatct ggcgaagtag cagaatattc tacaatgcg gcacaatcag ctgatgatgg 240
cttagatag agtaaaaag cattagaaga aattgaagtt ttgtatgaaa aaatgcaaat 300
cgcttctgat ttggtggatt cactcacaca aagaagcaat gaaattacaa gcgtaatctc 360
gcttattgat gatattgagg aacaacaaa cctccttggc ctcaatgcag ctattgaagc 420
accgagagct ggggagcacg gacgaggctt tgccgtgggc gctgatgagg tgcgaaaact 480
c 481

```

40 <210> 36  
 <211> 503  
 <212> ADN

ES 2 405 405 T3

<213> Desconocido

<220>

5 <223> Descripción de secuencia desconocida: clon 77538 Cbir-73 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

<400> 36

```

agctgttgat ggcgacatgg acaaagctat tgatttcttg cgtgaaaaag gtatggcaaa 60
ggctgctaag aagagcgacc gtgttgctgc tgaagggtta gctgatggtg aagttgtagg 120
_gaacatagct_gcagttgtg aaatcaacgc tgaaacagac tttgttgctc aaaaccaaca 180
attcaaggac ctgtaaadac gtgttgacag tttgatcgca gaaaataaac cagctgactt 240
agaagctget ttggctatca agactgacaa aggtacgac aacgaagaaa tcatcgaagc 300
tacacaagtt atcggtgaaa agatcacatt gcgtcgtttt gaattagttg aaaaagctga 360
caatgaaaac ttcgggtgctt acctacacat ggggtggaag atcgccggtt taactgtggt 420
tgaaggtgct gatgaagtgg ctgctaaaga cgttgcaatg cacgttgcaag ctatcaatcc 480
taagtacgtg aaccgcgacc aag                                     503

```

10

<210> 37

<211> 587

<212> ADN

<213> Desconocido

15

<220>

<223> Descripción de secuencia desconocida: clon 77539 Cbir-78 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

20

<400> 37

```

cagtgtttct cgggataatc cttttttccc atttctctatt tcggataaaa aatttactga 60
aatgtcaacg ctttccgcaa attcggcctg cgtaaaatca ttcagatcc gcagctgccg 120
aatcggttgc ccaatttcgg ttttgtttaag agtatcttcc ataggttccct ccagaaaagc 180
ggctgcttgt aacagtgggg aaactacagg actgttacgg ccaaattgcc atgattatgg 240
aatgtttgtg ttgctattag tttattttac caatccttaa aaaatagaaa acaataaaaa 300
gattgaatta atttctctatt agcgataaaa tataaactaa aagtgacgat atacataata 360
agactggata gggaggagtg gttggaatgg tgattgcaaa taatctattg tcccagttca 420
cagcaaggca gttaaatata aatagcggta agaaagaaaa agcggcagag aaactttctt 480
ccggctaccg cattaacagg gcgtcagata atgcggcggg cttaaaaatt tcggaaaaaa 540
tgcgtatgca gatccgcggg cttatgcceg ggcgcacaaa ataccca 587

```

<210> 38

<211> 422

25

<212> PRT

<213> Desconocido

<220>

30 <223> Descripción de secuencia desconocida: clon 76779 Cbir-1 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

<400> 38

ES 2 405 405 T3

Gly Gly Ile Ile Met Val Val Gln His Asn Leu His Ala Met Asn Ser  
5 10 15

Asn Arg Met Leu Gly Ile Thr Gln Lys Thr Ala Ser Lys Ser Thr Glu  
20 25 30

Lys Leu Ser Ser Gly Tyr Ala Ile Asn Arg Ala Ala Asp Asn Ala Ala  
35 40 45

Gly Leu Ala Ile Ser Glu Lys Met Arg Lys Gln Ile Arg Gly Leu Thr  
50 55 60

Gln Ala Ser Thr Asn Ala Glu Asp Gly Ile Ser Ser Val Gln Thr Ala  
65 70 75 80

Glu Gly Ala Leu Thr Glu Val His Asp Met Leu Gln Arg Met Asn Glu  
85 90 95

Leu Ala Ile Gln Ala Ala Asn Gly Thr Asn Ser Glu Asp Asp Arg Ser  
100 105 110

Tyr Ile Gln Asp Glu Ile Asp Gln Leu Thr Gln Glu Ile Asp Arg Val  
115 120 125

Ala Glu Thr Thr Lys Phe Asn Glu Thr Tyr Leu Leu Lys Gly Asp Thr  
130 135 140

Lys Asn Val Asp Ala Met Asp Tyr Thr Tyr Ser Tyr Lys Ala Val Thr  
145 150 155 160

Thr Asn Thr Val Ala Arg Ala Ser Val Leu Ala Ala Glu Asn Thr Ala  
165 170 175

Thr Gly Met Ser Val Ser Ile Ser Phe Ala Ala Asn Ser Gly Lys Val  
180 185 190

Thr Ala Ala Asp Ser Asn Asn Leu Ala Lys Ala Ile Arg Asp Gln Gly  
195 200 205

Phe Thr Ile Thr Thr Ser Thr Gln Asn Gly Lys Val Val Tyr Gly Leu  
210 215 220

Glu Leu Asn Gly Ser Asp Ala Lys Ala Asn Tyr Thr Val Ser Thr Val  
225 230 235 240

Ser Met Glu Ala Gly Thr Phe Lys Ile Leu Asn Ser Asn Lys Gln Val  
245 250 255

Val Ala Ser Val Thr Ile Ser Thr Thr Ala Ser Phe Lys Lys Val Ser  
260 265 270

Gly Met Ser Gln Ile Val Thr Ala Tyr Ser Val Ser Ala Ala Tyr Ala  
275 280 285

---

ES 2 405 405 T3

Thr Gly Asp Val Tyr Ser Leu Tyr Asp Ala Asp Gly Asn Ala Ile Ser  
 290 295 300

Ala Asn Lys Leu Asp Lys Tyr Phe Thr Ala Gly Gly Ala Thr Glu Ala  
 305 310 315 320

Gly Gly Ile Ala Thr Thr Leu Ser Ala Asn Ser Gly Val Pro Lys Val  
 325 330 335

Tyr Asp Val Leu Gly Lys Glu Val Ser Ala Val Ser Ile Ala Ser Thr  
 340 345 350

Leu Val Thr Ala Val Lys Asp Lys Thr Ala Ala Leu Lys Met Asn Phe  
 355 360 365

His Val Gly Ala Asp Gly Thr Asp Asn Asn Lys Ile Lys Ile Asn Ile  
 370 375 380

Glu Ala Met Thr Ala Lys Ser Leu Gly Val Asn Gly Leu Lys Val Ser  
 385 390 395 400

Gly Ser Ser Gly Thr Asn Ala Thr Asn Ala Ile Glu Ile Ile Ala Gly  
 405 410 415

Ala Ile Lys Lys Val Ser  
 420

<210> 39  
 <211> 73  
 <212> PRT  
 5 <213> Desconocido

<220>  
 <223> Descripción de secuencia desconocida: clon 76780 Cbir-2 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

<400> 39

Met Leu Val Asp Asn Cys Ala Met Gln Leu Val Lys Asp Pro Arg Gln  
 5 10 15

Phe Asp Val Ile Leu Thr Glu Asn Met Phe Gly Asp Ile Leu Ser Asp  
 20 25 30

Glu Ala Ser Met Val Thr Gly Ser Ile Gly Met Leu Ser Ser Ala Ser  
 35 40 45

Leu Asn Asp Thr Lys Phe Gly Leu Tyr Glu Pro Ser Gly Gly Ser Ala  
 50 55 60

Pro Asp Ile Ala Gly Lys Gly Ile Ala  
 65 70

15 <210> 40  
 <211> 105  
 <212> PRT  
 <213> Desconocido

20 <220>  
 <223> Descripción de secuencia desconocida: clon 76959 Cbir-3 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

<400> 40

ES 2 405 405 T3

```

Met Val Gly Thr Leu Val Gly Leu Ile Asn Met Leu Lys Ala Met Asp
      5                                10                                15

Ile Glu Thr Val Gly Gly Asn Leu Gly Pro Ala Met Ala Thr Ala Leu
      20                                25                                30

Val Thr Thr Leu Tyr Gly Cys Val Leu Ala His Met Ile Phe Gly Pro
      35                                40                                45

Ile Ala Thr Gln Leu Arg Gln Arg Asp Glu Glu Glu Thr Leu Cys Lys
      50                                55                                60

Leu Ile Ile Val Glu Gly Leu Met Ser Ile Gln Ala Gly Ala Asn Pro
      65                                70                                75                                80

Lys Phe Leu Arg Glu Lys Leu Leu Thr Phe Val Thr Gln Lys Gln Arg
      85                                90                                95

Gly Glu Asn Gly Gly Lys Lys Gly Lys
      100                                105

```

5 <210> 41  
 <211> 76  
 <212> PRT  
 <213> Desconocido

10 <220>  
 <223> Descripción de secuencia desconocida: clon 76959 Cbir-3 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/HeJ Bir

<400> 41

```

Met Ala Ser Ile Lys Lys Lys Ser Ser Gly Gly Gly Gly Ala Asn Trp
      5                                10                                15

Met Asp Thr Tyr Gly Asp Met Val Thr Leu Leu Leu Cys Phe Phe Val
      20                                25                                30

Leu Pro Val Phe His Val His Asp Arg Leu Gly Glu Val Glu Asp Asp
      35                                40                                45

Arg Ser Arg Ala Ser Ile Arg Thr Gln Ser Ser Ala Thr Ile Ser Pro
      50                                55                                60

Pro Asp Arg Thr Ala Leu Lys Ala Ala Arg Ala Ala
      65                                70                                75

```

15 <210> 42  
 <211> 152  
 <212> PRT  
 <213> Desconocido

20 <220>  
 <223> Descripción de secuencia desconocida: clon 76781 Cbir-6 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/HeJ Bir

25 <400> 42



ES 2 405 405 T3

<220>

<223> Descripción de secuencia desconocida: clon 76967 Cbir-16 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

5 <400> 44

Met	Ser	Arg	Ile	Gly	Arg	Leu	Pro	Ile	Ala	Val	Pro	Ala	Gly	Val	Thr
				5					10					15	
Val	Glu	Ile	Ala	Glu	His	Asn	Val	Val	Thr	Val	Lys	Gly	Pro	Lys	Gly
			20					25					30		
Thr	Leu	Val	Arg	Glu	Leu	Pro	Val	Glu	Met	Glu	Ile	Lys	Gln	Glu	Gly
		35					40					45			
Glu	Glu	Ile	Val	Val	Thr	Arg	Pro	Asn	Asp	Leu	Lys	Arg	Met	Lys	Ser
	50					55					60				
Leu	His	Gly	Leu	Thr	Arg	Thr	Leu	Ile	Asn	Asn	Met	Val	Ile	Gly	Val
	65				70					75					80
Ser	Gln	Gly	Tyr	Glu	Lys	Val	Leu	Glu	Val	Asn	Gly	Val	Gly	Tyr	Arg
				85					90					95	
Ala	Ala	Lys	Ser	Gly	Asn	Lys	Leu	Thr	Leu	Ser	Leu	Gly	Tyr	Ser	His
			100					105					110		
Pro	Val	Glu	Met	Val	Asp	Pro	Glu	Gly	Ile	Glu	Thr	Val	Leu	Glu	Gly
		115					120					125			
Gln	Asn	Lys	Ile	Thr	Val	Lys	Gly	Ile	Asp	Lys	Glu	Lys	Val	Gly	
	130					135					140				

<210> 45

10 <211> 153

<212> PRT

<213> Desconocido

<220>

15 <223> Descripción de secuencia desconocida: clon 76969 Cbir-20 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

<400> 45

Val	Tyr	Met	Ile	Ile	Gln	His	Asn	Ile	Ala	Ala	Ile	Asn	Ser	Tyr	Arg
				5					10					15	
Asn	Leu	Gly	Val	Asn	Gln	Ser	Gly	Leu	Asn	Lys	Asn	Leu	Glu	Lys	Leu
			20					25					30		
Ser	Ser	Gly	Tyr	Lys	Ile	Asn	Arg	Ala	Gly	Asp	Asp	Ala	Ala	Gly	Leu
		35					40					45			
Ala	Ile	Ser	Glu	Ser	Met	Arg	Ser	Gln	Ile	Asn	Gly	Leu	Asn	Gln	Gly
	50					55					60				
Gly	Leu	Asn	Lys	Asn	Leu	Glu	Lys	Leu	Ser	Ser	Gly	Tyr	Lys	Ile	Asn
	65				70					75					80
Arg	Ala	Gly	Asp	Asp	Ala	Ala	Gly	Leu	Ala	Ile	Ser	Glu	Ser	Met	Arg
			85					90						95	
Ser	Gln	Ile	Asn	Gly	Leu	Asn	Gln	Ala	Val	Asn	Asn	Ala	Lys	Asp	Ala
			100					105					110		
Ile	Gly	Leu	Ile	Gln	Thr	Ala	Glu	Gly	Ala	Leu	Thr	Arg	Asp	Ser	Ala
	115						120					125			
Tyr	Pro	Phe	Pro	Ala	Val	Cys	Cys	Arg	Ile	Pro	Ser	Leu	Ile	Arg	Asn
	130					135					140				
Gly	Leu	Asn	Thr	Gln	Glu	Leu	Asn	Lys							
	145				150										

20

<210> 46

ES 2 405 405 T3

<211> 110  
 <212> PRT  
 <213> Desconocido

5 <220>  
 <223> Descripción de secuencia desconocida: clon 76972 Cbir-27 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

<400> 46

10

```

Met Met Cys Ile Arg Asn Leu Arg Thr Glu Arg Ser Ser Ile Trp Ile
                    5                      10                      15

Arg Phe Val Ile Pro Leu Lys Arg Glu Gln Ile Arg Leu Ser Ser Gly
                    20                      25                      30

Lys Thr Gln Gly Lys Ala Gly Asp Pro Met Pro Tyr Ile Pro Pro Glu
                    35                      40                      45

Val Ile Glu Gln Ala Arg Gln Ile Asp Leu Leu Ser Tyr Met Lys Ala
                    50                      55                      60

Phe Glu Pro Asn Glu Leu Val Arg Ile Ser Gly Asn Asn Tyr Thr Thr
65                      70                      75                      80

Arg Thr His Asp Ser Leu Lys Ile Ser Asn Gly Lys Trp Met Trp Trp
                    85                      90                      95

Ser Gln Arg Ile Gly Gly Tyr Asn Ala Leu Asp Tyr Leu Ile
                    100                      105                      110
  
```

15 <210> 47  
 <211> 22  
 <212> PRT  
 <213> Desconocido

20 <220>  
 <223> Descripción de secuencia desconocida: clon 76974 Cbir-32 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

<400> 47

25

```

Met Gly Val Phe Gln His Ala Val Phe His His Val Leu Gly Ala Gln
                    5                      10                      15

Gln Leu Phe Leu Gly Gly
                    20
  
```

30 <210> 48  
 <211> 140  
 <212> PRT  
 <213> Desconocido

35 <220>  
 <223> Descripción de secuencia desconocida: clon 76975 Cbir-39 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

<400> 48

ES 2 405 405 T3

```

Asp Asn Thr Lys Lys Arg Gln Gly Asn Gly Lys Gln Gly Ser Tyr Met
                    5                      10                      15
Arg Ile Gln His Asn Ile Ser Ala Ile Asn Ser Gln Arg Asn Ile Met
                    20                      25                      30
Thr Thr Asp Arg Ala Leu Ser Lys Asn Leu Glu Lys Leu Ser Ser Gly
                    35                      40                      45
Tyr Arg Ile Asn Arg Ala Gly Asp Asp Ala Ala Gly Leu Ala Ile Ser
                    50                      55                      60
Glu Ala Met Arg Asn Gln Ile Thr Ala Met Asn Gln Gly His Glu Glu
                    65                      70                      75
Cys Ala Gly Arg Asn Phe Pro Gly Ala Asp Gly Arg Arg Arg Pro Asp
                    85                      90                      95
Arg Ser Pro Tyr His Ala Gln Pro His Glu Gly His Gly Gly Ala Gly
                    100                     105                     110
-----
Arg Gln Arg Asp Ile His Gly Ile Arg Thr Gly His Val Lys Leu Arg
                    115                      120                      125
Asp Gly Arg Ala Lys Gly Arg Asp Tyr Thr Tyr Trp
                    130                      135                      140

```

5 <210> 49  
 <211> 198  
 <212> PRT  
 <213> Desconocido

10 <220>  
 <223> Descripción de secuencia desconocida: clon 76977 Cbir-44 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/HeJ Bir

<400> 49

15

```

Ala Ala Glu Arg Leu Ser Ser Gly Tyr Arg Val Asn Arg Ala Ala Asp
                    5                      10                      15
Asp Ala Ala Ala Met Ala Ile Ser Glu Lys Lys Arg Ala Gln Ile Arg
                    20                      25                      30
Gly Leu Ala Arg Ala Ser Lys Asn Ala Gln Asp Gly Ile Ser Phe Val
                    35                      40                      45
Gln Thr Gly Asp Gly Ala Met Ser Gln Ile Gly Ala Met Leu His Arg
                    50                      55                      60
Met Arg Glu Leu Thr Val Gln Ala Leu Asn Asp Gly Val Tyr Glu Pro
                    65                      70                      75
Ala Asp Arg Ala Ala Leu Gln Met Glu Phe Asp Gln Leu Gln Gly Glu
                    85                      90                      95
Ile Asp Arg Val Asn Asp Gln Thr Glu Phe Asn Lys Lys Pro Val Phe
                    100                     105                     110
Glu His Tyr Thr Asp Asn Phe Ser Leu Leu Glu Gly Asn Arg Val Trp
                    115                      120                      125
Ser Gln Asp Gln Ile His Thr Ile Asp Ser Ser Asn Ser Ser Leu Thr
                    130                      135                      140
Val Lys Tyr Ile Ala Val Glu Pro Asp Gly Thr Glu Val Glu Lys Glu
                    145                      150                      155                      160
Lys Thr Leu Thr Ile Pro Glu Gly Thr Tyr Thr Thr Arg Glu Leu Met
                    165                      170                      175
Asp Glu Met Asp Asn Val Val Ser Ala Leu Gly Asp Glu Ala Asp Gly
                    180                      185                      190
Leu Tyr Leu Glu Tyr Ser
                    195

```

<210> 50



ES 2 405 405 T3

```

cctggcagag cacgccgcac gccataaccg ggagcggcat cgcattcaaa tcctccgcca 60
actggaaata ctccatataa ccaagcccca gcgtcatcat ataaccccag acattaaaat 120
tttccctgcg tttctcgacg atatccaccg aatccttcca atcataaaca ttatcccaga 180
tataggaacc ttccgagata cagccgccag gaaagcgaag gaaagtagga tgcagctctc 240
tcattgtctc tacaaggctc cggcgcagac ggttaattcg attcctaaga taatttttat 300
gggcgctcgc gctgccctcc tcctcaccaa atccccagac atgttccggg aacatcgaaa 360
ccatategat ggatatactg ccgaaaaaag ccaattctaa ctgcccgaag cactccttct 420
ctgccgtaag taccaccgct tgtttttccc cgtatttctt ccagtcgccg gatatggaaa 480
ttttttccac attgctgacc gccgccccgc a                                     511

```

<210> 53  
 <211> 213  
 5 <212> ADN  
 <213> Desconocido

<220>  
 <223> Descripción de secuencia desconocida: clon 73263 CB1-T4 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/HeJ Bir

<220>  
 <221> base modificada  
 <222> (1)..(213)  
 15 <223> n = g, a, c or t

<400> 53

```

atcgggaaaa agaaatactt tttatggatt tgcgacagat aggaagtaca tatgaaaaga 60
gtaccctttt ccgtatttct gccgcccggg tcggtcagaa atggcttgtt gggaagtgtc 120
ccaaaccctc tatgagaacg gaaaggagcg aaaacccatg aacggacgca aaaggacggn 180
acaggtcaaa ttctatgnga cagangaaga aac                                     213

```

<210> 54  
 <211> 511  
 <212> ADN  
 <213> Desconocido

<220>  
 <223> Descripción de secuencia desconocida: clon 73264 CB1-T5 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/HeJ Bir

<400> 54

```

gtacagggtc tgggtctcga ccccgcagtg ggccagaagc cgggaggccc cgttgccctt 60
ggtgcgcagc aggccagga gcagatgctc ggagttcacc gcgttctggc ccagccggcg 120
gctctcctcc accgcgccct ggatggctga gcagcagttg ggcgtaagtc cctgaaagct 180
ggactgggcy ggcactcccg tccccaccg ctgggcgatg gcggagcgca gggcctggct 240
gtccgcccc gcccgcgca gggccatggc ggcggggctg aactcctggc cggccagccc 300
aagcagaaga tgctcgtgca atagccaaat aaaacccatg cactaatggt atacaatctt 360
cctgaaaaac agccattact gcatgtaatc aatctccctt ttgcttttgt ttcaatatga 420
agcgggtttt cccacaatat ctgcattaaa taagataaat taccctcccc atatattctc 480
tgatcgggtc gtgtttttta tatatattca c                                     511

```

<210> 55  
 <211> 511  
 <212> ADN  
 <213> Desconocido

<220>  
 <223> Descripción de secuencia desconocida: clon 73266 CB1-T7 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/HeJ Bir

<400> 55

ES 2 405 405 T3

```

cctacaggcc aggacatgac ggcttgccg agctttacct gaattgccgc cgccttgtat 60
atgcctacaa taatctccat cctgacaaga ttggggaaag ggacgacctg attcgaagga 120
ttctcggcaa atgcggcgag cgtgtggcag tggagccgcc ctttcattgt gattatggcg 180
ttcatattga agttggcgac aatTTTTTTT ccaacttcaa ttgcgtaatc ctggatgttg 240
ccagggtcat gatcggcaaa aatgtcatgt ttgcgcccaa tgtgggaatt tatgcggctg 300
gccatcctgt tcattggcaa agccgcaatt cgggctatga atatggtcgg gaaatcagga 360
ttggcgacaa tgtctggctt ggcgggaatg ttatcgtcaa tcccgggata aatataggca 420
ataatgttgt tgtaggctcg ggcagcgttg ttaccaggga tatcccggac aatatgctgg 480
cagcgggtaa tccctgcagg gttttgccgc a 511

```

<210> 56  
 <211> 279  
 5 <212> ADN  
 <213> Desconocido

<220>  
 <223> Descripción de secuencia desconocida: clon 73267 CB1-T8 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/HeJ Bir

<400> 56

```

gataattatg aggaggcctt ggacagtgtg gaggaagtca agcgttccct tctggtagcg 60
ttggtggacc gcaaggtgag caaatatttc tccgagcggg acagcattat taagaaaata 120
gaaaaagaca aatattttgt ggcatttaag caaaaatatac tggacgagct gattgaggca 180
aagtttagta tccttgagga cgtcaagact gttaaagtgg gaaatgagat ggcagtcacc 240
ttgagcattg ggggtggcgt cagcgggaac agttatacc 279

```

15 <210> 57  
 <211> 234  
 <212> ADN  
 <213> Desconocido

<220>  
 <223> Descripción de secuencia desconocida: clon 73268 CB1-T9 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/HeJ Bir

25 <400> 57

```

cacttacgaa gatgtggtac tgccgcctct atggcatcat aatattacag tatttgatgc 60
ctagagcatg tctgaaaatt cattcccgcc atcttcaacgc cccactttgc gatttatttg 120
gctctcattc ggtcaacgta gctcactacg cctaccacat gatctacttt cgaaaaaaca 180
accttaata accttacatt tcaccagtct atttttctgg cgaaatgtgt ggtc 234

```

30 <210> 58  
 <211> 164  
 <212> ADN  
 <213> Desconocido

<220>  
 <223> Descripción de secuencia desconocida: clon 73269 CB1-T10 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/HeJ Bir

<400> 58

```

acattgcgaa atacaatctc acctttcggc tggaccaact cctgcgcgctc cggctcatca 60
aaaatcteta tadcagcadc cataatctgc aagaaccgct caatgccagt gattccgcgc 120
tgaaactgct ctgcaaaactc aataatcctg cggattgtcg caac 164

```

45 <210> 59  
 <211> 511  
 <212> ADN  
 <213> Desconocido

<220>  
 <223> Descripción de secuencia desconocida: clon 73270 CB1-T11 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/HeJ Bir

ES 2 405 405 T3

<220>  
 <221> base modificada  
 <222> (403)  
 5 <223> n = g, a, c or t

<400> 59

```

gagaaaaaat tttctttgaa ccgaaaaaaa caaaggctgc cattgtgaca tgcggcggtt 60
tgtgtccggg gctcaacgat gtgatccgtt ccattgtcat ggaattatgg catggatacg 120
gcgtgcgggg cattatcggc attccttacg gcttggaggg gtttatccct aaaaaatcacg 180
gacataagct tattgagctg acccctgatg ctgtttcaac aatttatatg ttgggcggaa 240
caattttaag ttcttcacgc ggagcgcaga attttgacga aatagcggat tttttgaatg 300
agaaagggat taatttgctt ttcgtgcttg gcggcgacgg taccatgaaa gccgcaaagg 360
ctatttcaac agcggtgaaa gagcatggtt tgaccctgtc ggntatcggg attcctaaaa 420
ctattgataa tgacatcaat tttgtggagc agtctttcgg gttcgccacc gctggtgacg 480
agggcagaaa agccattgcg gcggcgcata c                                     511
  
```

10 <210> 60  
 <211> 208  
 <212> ADN  
 <213> Desconocido

15 <220>  
 <223> Descripción de secuencia desconocida: clon 73272 CB1-T13 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/HeJ Bir

20 <220>  
 <221> base modificada  
 <222> (1)..(208)  
 <223> n = g, a, c or t

25 <400> 60

```

actcggcgta ccgggccatg gtatgggggt actggtactg ggagaagggt cccatcttgg 60
gcgggcactc cgcgctggtg tattccacca ccagggagat cagcagtgcg ttggctacgc 120
cgtggggcag gtggtggaac gcgccagat tgnngggccat ggagtggcat acncccatca 180
accaattngc aaaggccatg ccagccag                                     208
  
```

30 <210> 61  
 <211> 275  
 <212> ADN  
 <213> Desconocido

35 <220>  
 <223> Descripción de secuencia desconocida: clon 73273 CB1-T14 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/HeJ Bir

<400> 61

```

agcgtggaaa tcccatccag tgtgaaatgg atattggatt ctgccttttc aagatgcagc 60
agcttaacaa gcgtggcaat tcccccaagc gtgacagggg tagcttataa tgctttctta 120
ggctgcagcg acaacctcgt aatcgttggg actcctgggt cagaagcaga gaaatgca 180
caacagaaca acattacatt taaatttctt gattccgttc aggacatc cgaagcatct 240
ataaccttgg agagggcaag ctacacctat gacgc                                     275
  
```

40 <210> 62  
 <211> 357  
 <212> ADN  
 45 <213> Desconocido

<220>  
 <223> Descripción de secuencia desconocida: clon 73274 CB1-T15 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/HeJ Bir

50 <400> 62

ES 2 405 405 T3

```
cctgtttctt taaacatggc gtccagccgt ttctgaatat tctcccagat ggcatagccg 60
ttgggccgga taatcataka tccctttata ttggaatact ccaccagctc cgccttgcgc 120
accacatccg tataccactg cgcgaaatcc tctgttcacg acgtaatcgc ctcaaccaac 180
ttcttttctt ttgccataat aacctccatt caaatttgca gcttccgctc caggcacttc 240
tgaaccggga gaattttcag tcattcaaat ctctgtaaat cctgccgggg gttttacggt 300
ccgctgctgt tttatgatag ttccgcggtt tccttcccat tgccgcacgc acttcgc 357
```

<210> 63  
 5 <211> 376  
 <212> ADN  
 <213> Desconocido

<220>  
 10 <223> Descripción de secuencia desconocida: clon 73275 CB1-T16 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

<400> 63

```
tatacgtcca aagacagtat tgacattctg gagaagttct ttccggacga taaggaagcc 60
agggagaaaa aataagtagc tataagatat ttttaggtga tcaggccgga ttggaggaaa 120
tatatgataa cattgaccga gaatgaaaag aggatgggat ttccagctgga aggctgcaac 180
agatatgatg cgatacagga gattgctgtg ttctgccagt ataccctgga ccatgacaaa 240
aggaacattg ccgaaaatct gttaaagaaa ctctcgtgac tctcagaaaa tgactgtagg 300
gagctgattt gtgatattca gaaaaattat aatctgcctc ggaaagcaaa gactgttggg 360
15 gagatgattg ctgtgc 376
```

<210> 64  
 <211> 626  
 <212> ADN  
 20 <213> Homo sapiens

<400> 64

```
tgtagatctc ctccgccgcg cacagctcca ccagccggcc caggtacatc acccccacca 60
cgtcggagac atagcgcacc atggacaggt cgtgggcat gaacagatag gccagcccc 120
gctcctgctg gtagtcccgg aggaggttga ccacctgggc ctggatggac acatccagag 180
cggagatggg ctctgctcag atgaccagat cgggctggag gatcagcgcc cgggcatgct 240
ccaccctctg gcgctggccg ccggagaact cgtgggctga gcgcccggcg tgctccgccc 300
tcagccccac ccgctccagc atggggtaga cgtagtctgt ggccctctgc ttggctctga 360
ccacatggtg ggccagcagg ggctcggcga tgatgtcccg gacggtcatc cgggcatgta 420
gggagggcga ggggtcctgg aagatcatct gcatcttgcg gcgcccgggc tggagctcct 480
tttgggacat gcgggtgatg tctcgcctg ccagcacgat ctcccggcg gtggggtcgt 540
acatgctgat gatagaccgg gcgcagctgg acttgccgca gcccgactcg cccaccaggc 600
25 ccaggtctc cccccggcg atggag 626
```

<210> 65  
 <211> 457  
 <212> ADN  
 <213> Desconocido

<220>  
 30 <223> Descripción de secuencia desconocida: clon 75038 CB3-T2 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

<400> 65

```
tcagacttcc agaatctcct gttttgcat attcccggtt atcgtcagtc aacgatgccg 60
ccgcatcgc tctttcccc ttgcaggctg aaaaatattc gctctgacag cctgcttact 120
ggccggattg agcgcagag gaatcagga ccaaaaaaag cagtgaaaa attttcacag 180
aaaatttgca cgtactttt caaattttcc acatccgttt ttctactgct ttttgatacc 240
ttaatcctgc atgagaacaa tagtatattg ataaatttaa ggctggact gctgttccag 300
aatctgttcc acattttccg ccgtaacaa gcgatagggg agactcacat agatgtcatc 360
caccagcgtc atatcctccg gcagctctgt gccctgcccc agagaatacg caatttccag 420
35 aatactgtgc gcctgtccgg gagcgtcgtt cagcaact 457
```

<210> 66  
 40 <211> 364

ES 2 405 405 T3

<212> ADN  
<213> Desconocido

<220>

5 <223> Descripción de secuencia desconocida: clon 75039 CB3-T3 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

<400> 66

```

cggcataggc gagttttggt tggattaca ttcgggcaag tctgcgaca agtcggaaat 60
tataagaaat atagaaacca cgctttcgct caccgccgaa aacgacggg aagagtttat 120
aaatgctggc gcgagaatac gcgaaacgcg cgaggcgctg cacgcgcccc ttgctggctt 180
gcataaaaag cgtaggctgg gctgttccgt ttacgagggg atagtttact atttgcaaaa 240
taaateccgcg cccgagcttt taaatataga aactactttc tacgattcct taaccaaaca 300
aaagttagag gattacgaaa atatgcttat taccgctcaa gcggcggcaa aggagtgcgg 360
cgcc
10

```

<210> 67

<211> 616

<212> ADN

15 <213> Desconocido

<220>

20 <223> Descripción de secuencia desconocida: clon 75040 CB3-T4 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

<400> 67

```

ggttcctgct gcttgacgca ttctcttcg cgagatcctc ccctcgcac cacgcatggc 60
tgtgctccac gaaggtgtac cggcgctacg tcgcccggtt caagcaagcc ggaggcattc 120
cgctttccac gaagatccg atagtcggcg tctcgtacgc catgatgggc gtcagcgcgc 180
tggtcgtcca gaagccgctg gtgtgggctg tcctgggggtg cgtggccgctc ttctcctgt 240
atctgatggc cgtccgcatt ccgaccatcg agcagaagcg tgtcgatcga gcccgggcag 300
aagacgtggc ctgagtcacg gcaagccacc aagagcggcg gcgccgcagc catggcgcac 360
atggctcacc agaaggtcga agccatggag gcgacgagac cccccgccta gtgctggcgg 420
agcagctatc agaaccgtgt cggagctggg cggtagaccg catcttcgac gagccgctgg 480
ccacgatgct gcctgcctc catcatggcc aagcttcgct tcaagcatgc tgagcgcctg 540
ttttctcgca ccggacattt cgaaacgcgc tggatgcaag gatgcaacca aagacgacac 600
gaacgtgaaa catacg
35

```

25 <210> 68

<211> 369

<212> ADN

<213> Desconocido

30 <220>

<223> Descripción de secuencia desconocida: clon 75041 CB3-T5 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

<400> 68

```

gctgtataaa aacggcgtgc ttgctgaaaa gaagcccgtg agctttttaca gcgggctgaa 60
cgtggttacg ctgccgcttt acaccgatga ggaggccacc ttcgattacg agttgaaggt 120
tgaggcggcg gagccggacg gagatttttc gccgcataac aacgtttgct attttacgca 180
gaaggtgacc aacgagcgca aggtgctggt tttggggcgc tctgccgcgg atttggcggc 240
gggcaaaaaga atttacggcg agaaggacgt aagctatata agcagcgtga agcaggtgcc 300
cgtgacggtg gaggatatgt gcggctatga cgaaattggt ttaagcaact tcgacgtgcg 360
cacggtaag
35

```

<210> 69

<211> 604

40 <212> ADN

<213> Desconocido

<220>

45 <223> Descripción de secuencia desconocida: clon 75042 CB3-T6 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

ES 2 405 405 T3

<400> 69

```

ggcctatgga gccgatggg tctattggga aagggaagc aaggacggt ccgtaaaaa 60
ggttggggaa cgctgccaga cctgtgcgga acggaatat caggacggt ctgatgaaaa 120
tgtttccttt aaggcggcag cccatatttc cccggaagca gcaggcagc cagtgcgcgc 180
ccatgaaggg gagcatgtgt ccaatgccta tacaaggcg gcaaagaat atgggaaggt 240
tgtgtcagca tccgtcagca tccatacttc tgtctgccg gaatgtggca ggacgtatgt 300
atcgggggga accacatcca cccgattaa ataccggct gacccttatg aaaagagcaa 360
aaaggtgctt ggggaggaag aagccaaag gaaaaatatt gattatgcag ctttaattaa 420
ggaggtaaga atgatgaaga aattcagaac gaaaaataac gcaatgatgt gcatgatgtt 480
ctgccttggt gcggcagggg tgctctaac ccctaatacg gcatttgcca agaaaattac 540
ccgggacagc caggccgaat ccatggcaag gaaaaagta aaaggcggga ctgtcgtaga 600
aatc 604

```

5

<210> 70  
<211> 282  
<212> ADN  
<213> Desconocido

10

<220>  
<223> Descripción de secuencia desconocida: clon 75044 CB3-T9 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

15 <400> 70

```

ggccacaata caaccatcac ccagactaat tcccgtagaa ctatttacac ccaaaaggca 60
atTTTTgCcg atactaattg gctcactatt tccgccactt aaaacaccaa gtatgctagc 120
cccaccgCcg acatcgctcc cctcacctac gacaacacta gagctaattgc gcccttcatt 180
catacaagca cccattgcac cagcattgaa attcacataa cttgctccgg gcattttgtgt 240
ataaccacct ttgccaaagt acgctccaaa acgggttttt cc 282

```

20 <210> 71  
<211> 544  
<212> ADN  
<213> Desconocido

25 <220>  
<223> Descripción de secuencia desconocida: clon 75045 CB3-T10 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

<400> 71

```

ggatttattc cttttatac tttgtggct tccaaatTTt gaatatagc tattttgaaa 60
ttattaagaa aatataaata catttttttg aaatTTttat ttaaaaatat attgacattt 120
agaaataaat atTTtataat acagttacct taatcaatct actttatcag cttcgcattt 180
tactgtctgg aggtaagctt gagcaacggt acaattgcga tttatgattc agaagcagac 240
tatgctttaa aacttgCaga atatTTtgcg ctgaaaaatg gattaaacta ttctgtctcg 300
gtattcaccg attataattc ccttaagaat tatttatctg aaaatgatat tgatatactt 360
ttaatttctg aagaatTTtg tatgtatata gaagaattac agaatgtatc caatctgttt 420
atccttacag atggaaatat agatgctgcg ctaaagcagt atgcatccct atacaaatat 480
cagccagccg ataggatgct tcgggatatc atgtcctggt atgcactctc ttcttccaga 540
gaaa 544

```

30

<210> 72  
<211> 664  
<212> ADN  
<213> Desconocido

35 <220>  
<223> Descripción de secuencia desconocida: clon 75046 CB3-T12 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

40

<400> 72

ES 2 405 405 T3

```

gatactatat cttgtatact caaagcttat tgcgacagcc cctaggctga cctgttacat 60
acaagcaaca ggtcatacgc cacttgtgca gcatacttcc tctgcacggg tctatgcata 120
aatcaaaaat cgagtaggag atatttcttc ccctactcgc cgccggattc tccatttact 180
taaatcgtac agctgttccg tatgccataa ctctcggtgc gccttgcatt accgcagccg 240
aagcataacg cacatttaa accgcatctg caccaagctg ctctgcttcc tcaaccatcc 300
gctttgttgc aagcgcctc gcttcattca tcataccgt atacgctttt aactcgccgc 360
caacaagcgt tttaaaactt tgtgtaatat ctttaccaaa gtttttactt tgaattgtac 420
tgccctttac caaaccaagc atttctgttt ctttctctga tatataatca gtatttacta 480
agatcatctt cttccccact ccttatttca ttcatctttt ccacaagcac atatacaaca 540
acgccgatta gcgcaagtgg tactactcca aacaaaattt tcgttatcaa caacactgac 600
atactaaaac atatcgccgc ataaaaacia aaataaatca ctggaataat agaaatcaca 660
attg

```

<210> 73  
 <211> 613  
 5 <212> ADN  
 <213> Desconocido

<220>  
 <223> Descripción de secuencia desconocida: clon 75047 CB3-T13 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

<400> 73

```

ggccaggccc gttatagga atatctccat gaatctcata atccggataa tgaaccgcat 60
attcttttaa aatttcaatg gccttgtccg cataggcctg ctctcctgta gccagccaaa 120
tactgccat agagaatgcc cttcataat tccccatt gatcagtcce caccaggcac 180
tgtcataggg ttctcctgta aatacctttt tacaggaagg gcatctgtgt ttatggggac 240
tgtttctgtc aaaatcaagc tttaccgagc agtccggaca atagtaataa agcgtccagt 300
tgccgatccc cctttcggga accattatct caccattaaa gatctcatct acttctcttt 360
ttagccgctt aattgtggca ggataacggc tgcttctctc tctcagatgt tttctctctt 420
catccgtaaa ctgaatcatt ttcttctcc ccttttatca caatttttcc tgtttctgat 480
acagccaaaag agccccttc ctttaaagc tgcctggctt caaagtctt ttaaactctg 540
taggcgtaca cccgtagaac tttttaagc tgctggaaaa ataccgctga ttgtcatacc 600
cgcagcgttc tgc

```

<210> 74  
 <211> 514  
 <212> ADN  
 20 <213> Desconocido

<220>  
 <223> Descripción de secuencia desconocida: clon 75048 CB3-T14 de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

<400> 74

```

ttcatcactt aatgaataaa gcactggcag gtccattgga aggaaacggc gtatctctgt 60
tttacagtca aattggtgta atgccttgtt ggcaacactt gaaagatatt ccatacgtg 120
gtattcatat ggctgtaatt cttctaatac aaagtccatg ctttcttctt ggaatggcgc 180
aagcggaagt gaatacatat gtgcaagttt ctttatcagt tcttcatcat atacataccc 240
tgcacatatt aaaagccttc cctgtgccat atataaaggc ctttaactgt taaatcttga 300
gactgagcca atccagtctg cttcaccac tttctttaa agctctcctg tgtaagtgcc 360
ctcgcttgtt tcaaatggca gataatctat aaacaagtgg aaaagcctgt catcccatac 420
tgacattgat ttaataagct tgttttatc tgtaataacc tctctgtctg caagcatttc 480
ccccctgtca gtcattctgc aataatccaa atca

```

<210> 75  
 <211> 1377  
 <212> ADN  
 <213> Desconocido

<220>  
 <223> Descripción de secuencia desconocida: flagelina X Cbir-1 de longitud completa de amplificación por PCR de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/Hej Bir

<400> 75

ES 2 405 405 T3

gtagtacagc	acaatcttag	agcaatgaat	tctaacagaa	tgtaggcat	cacacagga	60
tctttaaaca	aatctacaga	gaagctctca	tcaggctaca	aggtaaacag	ggcagcagat	120
gatgcagcgg	gtctttcaat	ttccgagaaa	atgagaaaac	agatcagagg	actgtcacag	180
gcatctttga	atgctgagga	tggtatcagt	gcagtgcaga	ccgcagaggg	cgcattgaca	240
gaagttcatg	acatggtgca	gagaatgaac	gagctggcag	taaaggctgc	aaacggcaca	300
aactctacat	cagaccgtca	gacaattcag	gacgaggtag	accagctcct	cacagaaatc	360
gaccgtgtcg	cagagaccac	caaattcaat	gagctgtata	cattgaaggg	tgatgaggac	420
aaggtgacaa	gatatctttc	agcacatgac	gcaggatag	aaggaacctt	gacacagggc	480
gctacaaacg	cgacattttc	aatggaccag	ttaaagtgtg	gcgacacat	catgatcgca	540
ggcagagagt	accatatcag	cggaaccaag	gcagagcagg	cagcaatcat	tacggcttct	600
gtgaagattg	gacagcaggt	tacgattgat	ggaatcatgt	atacctgttc	atcagtaagc	660
aatgctgaca	aatttgaact	gaaaagtgag	gatttgattg	caaaactcga	cacttcaagc	720
ctgagtatta	tgtcagtgaa	tggcaagacc	tactacggcg	caggcatcac	agatgacagg	780
actgttgtaa	gttcaattgg	tgcatacaag	ctgattcaga	aggagctcgg	actggcaagc	840
agcattggtg	cagacggcgc	aacacaggct	tcggtaaatg	ccggagtaga	tggcaagact	900
ttgatgaagc	cgagttttga	gggcaaatgg	gtatttagta	tcgacaaggg	aagcgttcag	960
gtacgcgagg	acattgattt	cagcctccat	gtaggtgcag	atgccgacat	gaacaacaag	1020
attgcggtga	agatcggagc	gcttgacacg	aagggacttg	gtatccaagg	actgaatgta	1080
aaggatacga	cagggcgcagc	agcgacctac	gcgattgatt	cgattgcgga	cgcagtgcca	1140
agaatttctg	cgcagcgctc	tttactcggg	gcagtgacga	accggttaga	gcacacgatc	1200
aacaacttgg	ataacgttgt	agagaacaca	accgccgcag	agagccagat	ccgtgataca	1260
gacatggcga	cagagatggt	gaagtactct	aataacaacg	tacttgacac	ggcagggccag	1320
tcaatgtag	cacagtctaa	tcaggcaaat	caggggtgtac	ttcagctctt	acagtaa	1377

- 5 <210> 76
- <211> 450
- <212> ADN
- <213> Desconocido

- 10 <220>
- <223> Descripción de secuencia desconocida: extremo conservado amino-terminal de flagelina X truncada de clon CBir-1 de amplificación por PCR de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/HeJ Bir

- 15 <400> 76

gtagtacagc	acaatcttag	agcaatgaat	tctaacagaa	tgtaggcat	cacacagga	60
tctttaaaca	aatcgacaga	gaagctatca	tcaggctaca	aggtaaacag	ggcagcagat	120
gatgcagcgg	gtctttcaat	ttccgagaaa	atgagaaaac	agatcagagg	actgtcacag	180
gcatctttga	atgctgagga	tggtatcagt	gcagtgcaga	ccgcagaggg	cgcattgaca	240
gaagttcatg	acatggtgca	gagaatgaac	gagctggcag	taaaggctgc	aaacggcaca	300
aactctacat	cagaccgtca	gacaattcag	gacgaggtag	accagctcct	cacagaaatc	360
gaccgtgtcg	cagagaccac	caaattcaat	gagctgtata	cattgaaggg	tgatgaggac	420
aaggtgacaa	gatatctttc	agcacattaa				450

- 20 <210> 77
- <211> 990
- <212> ADN
- <213> Desconocido

- 25 <220>
- <223> Descripción de secuencia desconocida: extremo conservado amino-terminal de flagelina X truncada de clon CBir-1 más parte variable de amplificación por PCR de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/HeJ Bir

- 30 <400> 77

ES 2 405 405 T3

```

gtagtacagc acaatcttag agcaatgaat tctaacagaa tgtaggcat cacacagga 60
tctttaaaca aatctacaga gaagctctca tcaggctaca aggtaaacag ggcagcagat 120
gatgcagcgg gtctttcaat ttccgagaaa atgagaaaac agatcagagg actgtcacag 180
gcatctttga atgctgagga tggatcagt gcagtgacga ccgagaggg cgcattgaca 240
gaagttcatg acatgttga gagaatgaac gagctggcag taaaggctgc aaacggcaca 300
aactctacat cagaccgtca gacaattcag gacgaggtag accagctcct cacagaaatc 360
gaccgtgtcg cagagaccac caaattcaat gagctgtata cattgaaggg tgatgaggac 420
aaggtgacaa gatattcttc agcacatgac gcaggtatag aaggaacctt gacacagggc 480
gctacaaacg cgacattttc aatggaccag ttaaagtttg gcgacacccat catgatcgca 540
ggcagagagt accatatcag cgaaccaag gcagagcagg cagcaatcat tacggcttct 600
gtgaagattg gacagcagg taccattgat ggaatcatgt atacctgttc atcagtaagc 660
aatgctgaca aatttgaact gaaaagtgag gatttgattg caaaactcga cacttcaagc 720
ctgagtatta tgtcagtga tggcaagacc tactacggcg caggcatcac agatgacagg 780
actgttgtaa gttcaattgg tgcatacaag ctgattcaga aggagctcgg actggcaagc 840
agcattggtg cagacggcgc aacacaggct tcggtaaatt ccggagtaga tggcaagact 900
ttgatgaagc cgagttttga gggcaaatgg gtatttagta tcgacaaggg aagcgttcag 960
gtacgcgagg acattgatt cagcctctaa 990

```

<210> 78  
 <211> 399  
 5 <212> ADN  
 <213> Desconocido

<220>  
 10 <223> Descripción de secuencia desconocida: extremo conservado carboxilo-terminal de flagelina X truncada de clon CBir-1 de amplificación por PCR de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/HeJ Bir

<400> 78

```

ttcagcctcc atgtaggtgc agatgccgac atgaacaaca agattgcggt gaagatcgga 60
gcgcttgaca cgaagggact tggatccaa ggactgaatg taaaggatac gacaggcgca 120
gcagcgacct acgcgattga ttcgattgcg gacgcagtg caagaatttc tgcgcagcgc 180
tctttactcg gtgcagtgca gaaccggta gagcacacga tcaacaactt ggataacggt 240
gtagagaaca caaccgccgc agagagccag atccgtgata cagacatggc gacagagatg 300
gtgaagtact ctaataacaa cgtacttgca caggcaggcc agtcaatggt agcacagtct 360
15 aatcaggcaa atcagggtgt acttcagctc ttacagtaa 399

```

<210> 79  
 <211> 458  
 20 <212> PRT  
 <213> Desconocido

<220>  
 25 <223> Descripción de secuencia desconocida: flagelina X CBir-1 de longitud completa de amplificación por PCR de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/HeJ Bir  
 <400> 79



ES 2 405 405 T3

Gln Ala Ala Ile Ile Thr Ala Ser Val Lys Ile Gly Gln Gln Val Thr  
 195 200 205

Ile Asp Gly Ile Met Tyr Thr Cys Ser Ser Val Ser Asn Ala Asp Lys  
 210 215 220

Phe Glu Leu Lys Ser Glu Asp Leu Ile Ala Lys Leu Asp Thr Ser Ser  
 225 230 235 240

Leu Ser Ile Met Ser Val Asn Gly Lys Thr Tyr Tyr Gly Ala Gly Ile  
 245 250 255

Thr Asp Asp Arg Thr Val Val Ser Ser Ile Gly Ala Tyr Lys Leu Ile  
 260 265 270

Gln Lys Glu Leu Gly Leu Ala Ser Ser Ile Gly Ala Asp Gly Ala Thr  
 275 280 285

Gln Ala Ser Val Asn Ala Gly Val Asp Gly Lys Thr Leu Met Lys Pro  
 290 295 300

Ser Phe Glu Gly Lys Trp Val Phe Ser Ile Asp Lys Gly Ser Val Gln  
 305 310 315 320

Val Arg Glu Asp Ile Asp Phe Ser Leu His Val Gly Ala Asp Ala Asp  
 325 330 335

Met Asn Asn Lys Ile Ala Val Lys Ile Gly Ala Leu Asp Thr Lys Gly  
 340 345 350

Leu Gly Ile Gln Gly Leu Asn Val Lys Asp Thr Thr Gly Ala Ala Ala  
 355 360 365

Thr Tyr Ala Ile Asp Ser Ile Ala Asp Ala Val Ala Arg Ile Ser Ala  
 370 375 380

Gln Arg Ser Leu Leu Gly Ala Val Gln Asn Arg Leu Glu His Thr Ile  
 385 390 395 400

Asn Asn Leu Asp Asn Val Val Glu Asn Thr Thr Ala Ala Glu Ser Gln  
 405 410 415

Ile Arg Asp Thr Asp Met Ala Thr Glu Met Val Lys Tyr Ser Asn Asn  
 420 425 430

Asn Val Leu Ala Gln Ala Gly Gln Ser Met Leu Ala Gln Ser Asn Gln  
 435 440 445

Ala Asn Gln Gly Val Leu Gln Leu Leu Gln  
 450 455

<210> 80  
 <211> 149  
 5 <212> PRT  
 <213> Desconocido

<220>

10 <223> Descripción de secuencia desconocida: extremo conservado amino-terminal de flagelina X truncada de clon CBir-1 de amplificación por PCR de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/HeJ Bir

<400> 80



ES 2 405 405 T3

Ala Glu Asp Gly Ile Ser Ala Val Gln Thr Ala Glu Gly Ala Leu Thr  
65 70 75 80

Glu Val His Asp Met Leu Gln Arg Met Asn Glu Leu Ala Val Lys Ala  
85 90 95

Ala Asn Gly Thr Asn Ser Thr Ser Asp Arg Gln Thr Ile Gln Asp Glu  
100 105 110

Val Asp Gln Leu Leu Thr Glu Ile Asp Arg Val Ala Glu Thr Thr Lys  
115 120 125

Phe Asn Glu Leu Tyr Thr Leu Lys Gly Asp Glu Asp Lys Val Thr Arg  
130 135 140

Tyr Leu Ser Ala His Asp Ala Gly Ile Glu Gly Thr Leu Thr Gln Gly  
145 150 155 160

Ala Thr Asn Ala Thr Phe Ser Met Asp Gln Leu Lys Phe Gly Asp Thr  
165 170 175

Ile Met Ile Ala Gly Arg Glu Tyr His Ile Ser Gly Thr Lys Ala Glu  
180 185 190

Gln Ala Ala Ile Ile Thr Ala Ser Val Lys Ile Gly Gln Gln Val Thr  
195 200 205

Ile Asp Gly Ile Met Tyr Thr Cys Ser Ser Val Ser Asn Ala Asp Lys  
210 215 220

Phe Glu Leu Lys Ser Glu Asp Leu Ile Ala Lys Leu Asp Thr Ser Ser  
225 230 235 240

Leu Ser Ile Met Ser Val Asn Gly Lys Thr Tyr Tyr Gly Ala Gly Ile  
245 250 255

Thr Asp Asp Arg Thr Val Val Ser Ser Ile Gly Ala Tyr Lys Leu Ile  
260 265 270

Gln Lys Glu Leu Gly Leu Ala Ser Ser Ile Gly Ala Asp Gly Ala Thr  
275 280 285

Gln Ala Ser Val Asn Ala Gly Val Asp Gly Lys Thr Leu Met Lys Pro  
290 295 300

Ser Phe Glu Gly Lys Trp Val Phe Ser Ile Asp Lys Gly Ser Val Gln  
305 310 315 320

Val Arg Glu Asp Ile Asp Phe Ser Leu  
325

<210> 82  
<211> 129  
<212> PRT  
5 <213> Desconocido

<220>

10 <223> Descripción de secuencia desconocida: extremo conservado carboxilo-terminal de flagelina X truncada de clon CBir-1 de amplificación por PCR de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/HeJ Bir

<400> 82





ES 2 405 405 T3

Gly Val Leu Ala Asp Thr Ile Asn Lys Phe Ser Asp Thr Leu Gly Val  
 225 230 235 240

Arg Ala Tyr Ala Thr Val Leu Gly Thr Gly Gly Val Pro Val Gln Ser  
 245 250 255

Gly Thr Val His Gly Leu Val Val Asn Gly Thr Thr Ile Gly Thr Ile  
 260 265 270

Asn Asp Val Arg Lys Asn Asp Ala Asp Gly Arg Leu Ile Asn Ala Phe  
 275 280 285

Asn Ser Ile Lys Glu Arg Thr Gly Val Glu Ala Tyr Val Asp Ile Glu  
 290 295 300

Gly Arg Leu Asn Leu Arg Ser Leu Asp Gly Arg Ala Ile Ser Val His  
 305 310 315 320

Ala Glu Gly Lys Thr Gly Ala Val Leu Gly Gly Gly Ser Phe Ala Gly  
 325 330 335

Val Ser Gly Thr Asn His Ala Ile Val Gly Arg Ile Ser Leu Val Arg  
 340 345 350

Thr Asp Ala Arg Asp Ile Ile Val Ser Gly Thr Asn Phe Ser Ser Val  
 355 360 365

Gly Phe His Ser Ala Gln Gly Ile Ala Gln Tyr Thr Val Asn Leu Arg  
 370 375 380

Ser Val Arg Gly Asn Met Asp Ala Asn Ile Ala Ser Ala Ser Gly Ala  
 385 390 395 400

Asn Ala Asn Ala Ala Gln Ala Val Gln Asn Lys Asp Gly Ile Gly Ala  
 405 410 415

Gly Val Thr Ser Leu Arg Gly Ala Met Val Val Met Asp Met Ala Glu  
 420 425 430

Ser Ala Thr Arg Gln Leu Asp Lys Ile Arg Ala Asp Met Gly Ser Val  
 435 440 445

Gln Met Gln Leu Val Ala Thr Ile Asn Asn Ile Ser Ile Thr Gln Val  
 450 455 460

Asn Val Lys Ala Ala Glu Ser Gln Ile Arg Asp Val Asp Phe Ala Gln  
 465 470 475 480

Glu Ser Ala Thr Phe Ser Lys His Asn Ile Leu Ala Gln Ser Gly Ser  
 485 490 495

Phe Ala Met Ala Gln Ala Asn Ala Val Gln Gln Asn Val Leu Arg Leu  
 500 505 510

Leu Gln

<210> 85  
 <211> 1494  
 <212> ADN  
 <213> Desconocido

5

<220>  
 <223> Descripción de secuencia desconocida: flagelina CBir-1 de longitud completa de amplificación por PCR de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/HeJ Bir

10

<400> 85

ES 2 405 405 T3

```

ggaggtatta ttatggtagt acagcacaat ttacaggcaa tgaactctaa cagaatgtta 60
ggcatcacac agaagacagc atctaagtct acagaaaagt tatcttcagg ttacgcaatc 120
aaccgcgcag cagacaacgc agcaggctct gctatctctg agaagatgag aaagcagatc 180
agaggactta cacaggcttc taaaaatgct gaggacggca tcagctctgt acagacagca 240
gaaggcgctt tgacagaagt gcatgatatg cttcagagaa tgaacgagct ggcaattcag 300
gcagcaaactc gcacaaactc agaagatgac cgctcataca ttcaggacga aattgaccag 360
ctgacacagg aatcgatcg tgttgctgag acaacaaagt tcaatgagac atatctcttg 420
aagggtgaca caaagaacgt tgacgctatg gactatacat atagctataa ggcagttaca 480
acgaatactg tagcaagagc ttcgggttta gcagcagaga acacagctac aggtatgtca 540
gtagtattt catttgctgc aaacagcggc aaggttactg cagctgactc taacaacctt 600
gcaaaggcta tcagagatca gggcttcaca atcacaacat ctaccagaa tggtagggtt 660
gtttacggtc ttgagctgaa cggaagcgat gcaaaggcaa actatacagt ttcaacagta 720
agtatggaag ctggtacatt caagatcctg aattctaata agcaggttgt tgcatctgta 780
acaatatcta caacagctag ctttaaaaag gtatctggta tgtcacagat cgttacggcg 840
tactctgtat cagcagctta tgcgacgggt gatgtatact ctctctatga cgcagacgga 900
aatgcaattt cagcaaaaa gctggataag tactttacgg caggcggcgc tacagaggca 960
ggcggaatag ctactacact ttcagcaaac tctgggtgtc ctaaggttta tgacgtactc 1020
ggaaaagagg tttctgcagt aagcattgca agtactttag taacagcagt taaggataag 1080
acggctgcat tgaagatgaa cttccatgta ggtgctgacg gaacagataa caacaagatt 1140
aagatcaaca ttgaggctat gacagctaag agtcttgag ttaacggctc gaaggtgagc 1200
ggttcgcgag gaacaaacgc tacaacgct atcgagataa tcgctggcgc tatcaagaag 1260
gtttctacac agagatctgc tcttggtgag gttcagaaca gattagagca cacaatcaac 1320
aacttgata acatcggtga gaacacaaca gcagctgagt caggaatccg cgatacagat 1380
atgctacag agatggttaa gtactctaac gctaatatcc tttcacaggc aggtcagctc 1440
atgcttgcac agtctaacca gtctaaccag ggtgtacttc agctcttaca gtaa 1494

```

<210> 86

<211> 493

5 <212> PRT

<213> Desconocido

<220>

10 <223> Descripción de secuencia desconocida: flagelina CBir-1 de longitud completa de amplificación por PCR de biblioteca de expresión de cizalladura al azar de ADN genómico de bacterias cecales de ratón C3H/HeJ Bir

<400> 86

```

Met Val Val Gln His Asn Leu Gln Ala Met Asn Ser Asn Arg Met Leu
          5              10              15

Gly Ile Thr Gln Lys Thr Ala Ser Lys Ser Thr Glu Lys Leu Ser Ser
          20              25              30

Gly Tyr Ala Ile Asn Arg Ala Ala Asp Asn Ala Ala Gly Leu Ala Ile
          35              40              45

Ser Glu Lys Met Arg Lys Gln Ile Arg Gly Leu Thr Gln Ala Ser Thr
          50              55              60

Asn Ala Glu Asp Gly Ile Ser Ser Val Gln Thr Ala Glu Gly Ala Leu
          65              70              75              80

Thr Glu Val His Asp Met Leu Gln Arg Met Asn Glu Leu Ala Ile Gln
          85              90              95

```

15

ES 2 405 405 T3

Ala Ala Asn Gly Thr Asn Ser Glu Asp Asp Arg Ser Tyr Ile Gln Asp  
100 105 110

Glu Ile Asp Gln Leu Thr Gln Glu Ile Asp Arg Val Ala Glu Thr Thr  
115 120 125

Lys Phe Asn Glu Thr Tyr Leu Leu Lys Gly Asp Thr Lys Asn Val Asp  
130 135 140

Ala Met Asp Tyr Thr Tyr Ser Tyr Lys Ala Val Thr Thr Asn Thr Val  
145 150 155 160

Ala Arg Ala Ser Val Leu Ala Ala Glu Asn Thr Ala Thr Gly Met Ser  
165 170 175

Val Ser Ile Ser Phe Ala Ala Asn Ser Gly Lys Val Thr Ala Ala Asp  
180 185 190

Ser Asn Asn Leu Ala Lys Ala Ile Arg Asp Gln Gly Phe Thr Ile Thr  
195 200 205

Thr Ser Thr Gln Asn Gly Lys Val Val Tyr Gly Leu Glu Leu Asn Gly  
210 215 220

Ser Asp Ala Lys Ala Asn Tyr Thr Val Ser Thr Val Ser Met Glu Ala  
225 230 235 240

Gly Thr Phe Lys Ile Leu Asn Ser Asn Lys Gln Val Val Ala Ser Val  
245 250 255

Thr Ile Ser Thr Thr Ala Ser Phe Lys Lys Val Ser Gly Met Ser Gln  
260 265 270

Ile Val Thr Ala Tyr Ser Val Ser Ala Ala Tyr Ala Thr Gly Asp Val  
275 280 285

Tyr Ser Leu Tyr Asp Ala Asp Gly Asn Ala Ile Ser Ala Asn Lys Leu  
290 295 300

Asp Lys Tyr Phe Thr Ala Gly Gly Ala Thr Glu Ala Gly Gly Ile Ala  
305 310 315 320

Thr Thr Leu Ser Ala Asn Ser Gly Val Pro Lys Val Tyr Asp Val Leu  
325 330 335

Gly Lys Glu Val Ser Ala Val Ser Ile Ala Ser Thr Leu Val Thr Ala  
340 345 350

Val Lys Asp Lys Thr Ala Ala Leu Lys Met Asn Phe His Val Gly Ala  
355 360 365

Asp Gly Thr Asp Asn Asn Lys Ile Lys Ile Asn Ile Glu Ala Met Thr  
370 375 380

Ala Lys Ser Leu Gly Val Asn Gly Leu Lys Val Ser Gly Ser Ser Gly  
385 390 395 400

Thr Asn Ala Thr Asn Ala Ile Glu Ile Ile Ala Gly Ala Ile Lys Lys  
405 410 415

ES 2 405 405 T3

```

Val Ser Thr Gln Arg Ser Ala Leu Gly Ala Val Gln Asn Arg Leu Glu
      420                               425                       430

His Thr Ile Asn Asn Leu Asp Asn Ile Val Glu Asn Thr Thr Ala Ala
      435                               440                       445

Glu Ser Gly Ile Arg Asp Thr Asp Met Ala Thr Glu Met Val Lys Tyr
      450                               455                       460

Ser Asn Ala Asn Ile Leu Ser Gln Ala Gly Gln Ser Met Leu Ala Gln
      465                               470                       475                       480

Ser Asn Gln Ser Asn Gln Gly Val Leu Gln Leu Leu Gln
      485                               490

```

<210> 87

<211> 6

<212> PRT

5 <213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: etiqueta de purificación de seis histidinas

10 <400> 87

```

His His His His His His
  -1                               5

```

**REIVINDICACIONES**

1. Polinucleótido aislado que comprende una secuencia seleccionada de entre el grupo que consiste en:
- 5 (a) la secuencia proporcionada en SEC ID N° 75;
- (b) el complemento de la secuencia proporcionada en SEC ID N° 75;
- 10 (c) secuencias que presentan al menos el 75% de identidad con la secuencia de SEC ID N° 75; y
- (d) secuencias que presentan al menos el 90% de identidad con la secuencia de SEC ID N° 75.
2. Polipéptido aislado que comprende una secuencia de aminoácidos seleccionada de entre el grupo que consiste en:
- 15 (a) secuencias codificadas por un polinucleótido según la reivindicación 1;
- (b) la secuencia expuesta en SEC ID N° 79; y
- 20 (c) secuencias que presentan al menos el 70%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% o el 99% de identidad con la secuencia expuesta en SEC ID N° 79.
3. Fragmento de polipéptido aislado que comprende al menos 25, 50, 100 o más aminoácidos contiguos de la secuencia de SEC ID N° 79, siendo el fragmento de polipéptido inmunológicamente reactivo con un anticuerpo que se une al polipéptido de longitud completa de SEC ID N° 79.
- 25 4. Fragmento de polipéptido aislado según la reivindicación 3, en el que el fragmento de polipéptido comprende:
- (a) SEC ID N° 80, que corresponde a la región conservada amino-terminal de la secuencia expuesta en SEC ID N° 79; o
- 30 (b) SEC ID N° 81, que corresponde a la región conservada amino-terminal más la región variable de la secuencia expuesta en SEC ID N° 79.
- 35 5. Proteína de fusión que comprende un fragmento de polipéptido según la reivindicación 3 o la reivindicación 4.
6. Proteína de fusión que consiste en un fragmento de polipéptido según la reivindicación 4 y una etiqueta de seis histidinas (SEC ID N° 87).
- 40 7. Vector de expresión que comprende un polinucleótido según la reivindicación 1 unido funcionalmente a una secuencia de control de la expresión.
8. Célula huésped transformada o transfectada con un vector de expresión según la reivindicación 7.
- 45 9. Procedimiento de preparación de un polipéptido según la reivindicación 2, que comprende:
- cultivar una célula huésped transformada o transfectada con un vector de expresión según la reivindicación 7 en condiciones mediante las cuales se expresa el polinucleótido y se produce el polipéptido; y
- 50 recuperar el polipéptido.
10. Anticuerpo aislado, o fragmento de unión a antígeno del mismo, que se une específicamente al polipéptido de SEC ID 79, opcionalmente siendo dicho anticuerpo un anticuerpo monoclonal, por ejemplo un anticuerpo humano o un anticuerpo que se ha humanizado.
- 55 11. Anticuerpo según la reivindicación 10, en el que dicho anticuerpo se selecciona de entre el grupo que consiste en:
- (i) un anticuerpo neutralizante contra dicho polipéptido; y
- 60 (ii) un anticuerpo que bloquea la interacción entre dicho polipéptido y un receptor tipo Toll, opcionalmente en el que dicho receptor tipo Toll es TLR5.
- 65 12. Proteína de fusión, que comprende al menos un polipéptido según la reivindicación 2.
13. Procedimiento *ex vivo* de estimulación y/o expansión de células T específicas para una proteína bacteriana

entérica, por ejemplo una proteína de flagelina, que comprende poner en contacto células T con al menos un componente seleccionado de entre el grupo que consiste en:

- 5 (a) polipéptidos según la reivindicación 2;
- (b) polinucleótidos según la reivindicación 1; y
- (c) células presentadoras de antígeno que expresan un polinucleótido según la reivindicación 1,
- 10 en unas condiciones y durante un tiempo suficiente para permitir la estimulación y/o la expansión de células T.

14. Procedimiento según la reivindicación 13, en el que dichas células T son células T CD4+, en el que opcionalmente dichas células T CD4+ pueden mediar una disminución de una inflamación en el colon.

15 15. Composición, que comprende un primer componente seleccionado de entre el grupo que consiste en portadores fisiológicamente aceptables y un segundo componente seleccionado de entre el grupo que consiste en:

- (a) polipéptidos según la reivindicación 2
- 20 (b) polinucleótidos según la reivindicación 1;
- (c) anticuerpos según la reivindicación 10;
- (d) proteínas de fusión según la reivindicación 12;
- 25 (e) células presentadoras de antígeno según se define en la reivindicación 13 (c), y
- (f) células presentadoras de antígeno que expresan un polipéptido según la reivindicación 2.

30 16. Composición según la reivindicación 15, que comprende además un inmunoestimulante, en la que opcionalmente dicho inmunoestimulante es un adyuvante seleccionado de entre el grupo que consiste en un adyuvante incompleto de Freund; adyuvante completo de Freund; adyuvante 65 de Merck; AS-1, AS-2; gel de hidróxido de aluminio; fosfato de aluminio; una sal de calcio, hierro o zinc; una suspensión insoluble de tirosina acilada, azúcares acilados; polisacáridos derivados catiónica o aniónicamente; polifosfacenos; microesferas biodegradables; monofosforil lípido A, QS21, 4-fosfatos de aminoalquil-glucosaminida y Quil A.

35

17. Procedimiento de detección de la presencia de una enfermedad inflamatoria del intestino en un mamífero, que comprende las etapas siguientes:

- 40 (a) poner en contacto una muestra biológica obtenida de dicho mamífero que comprende anticuerpos con un polipéptido según la reivindicación 2;
- (b) detectar en la muestra una cantidad de anticuerpo que se une al polipéptido; y
- 45 (c) comparar la cantidad de anticuerpo unido con un valor de corte predeterminado y a partir de ello, determinar la presencia de una enfermedad inflamatoria del intestino en el mamífero.

18. Procedimiento según la reivindicación 17, en el que dicho polipéptido comprende un polipéptido de flagelina.

50 19. Procedimiento para diagnosticar una enfermedad inflamatoria del intestino (EII), que comprende las etapas siguientes:

- (a) poner en contacto una muestra biológica obtenida de un mamífero con un polipéptido según la reivindicación 2;
- 55 (b) detectar en la muestra una cantidad de anticuerpo que se une al polipéptido; y
- (c) comparar la cantidad de anticuerpo unido con un valor de corte predeterminado y a partir de ello diagnosticar que el mamífero presenta EII.
- 60

20. Procedimiento *ex vivo* de estimulación y/o expansión de células B que producen anticuerpos específicos para una proteína bacteriana entérica, que comprende poner en contacto células B con al menos un componente seleccionado de entre el grupo que consiste en:

- 65 (a) polipéptidos según la reivindicación 2; y

(b) polinucleótidos según la reivindicación 1;

en unas condiciones y durante un tiempo suficiente para permitir la estimulación y/o la expansión de células B.

5 21. Procedimiento según la reivindicación 20, en el que dichas células B producen anticuerpos seleccionados de entre el grupo que consiste en:

(i) anticuerpos que se unen a una proteína de flagelina, en el que opcionalmente dichos anticuerpos son anticuerpos neutralizantes contra dicha proteína de flagelina; y

10 (ii) anticuerpos que bloquean la interacción entre una proteína de flagelina y un receptor tipo Toll, en el que opcionalmente dicho receptor tipo Toll comprende TLR5.

15 22. Composición según la reivindicación 15 o la reivindicación 16, para su utilización en el tratamiento de una inflamación gastrointestinal asociada con una enfermedad inflamatoria del intestino en un mamífero, pudiendo dicha composición estimular una respuesta inmunitaria en dicho mamífero.

20 23. Composición según la reivindicación 22, en la que dicha respuesta inmunitaria comprende células T que producen una citocina seleccionada de entre el grupo que consiste en interleucina 10 (IL-10), interleucina 4 (IL-4), interleucina 5 (IL-5) y factor de crecimiento transformante beta (TGF $\beta$ ), produciendo preferentemente dichas células T IL-10 y/o TGF $\beta$ .

24. Procedimiento de identificación de antígenos bacterianos asociados con una enfermedad inflamatoria del intestino en un mamífero, que comprende las etapas siguientes:

25 (a) poner en contacto una muestra biológica obtenida de dicho mamífero que comprende células T con al menos un componente seleccionado de entre el grupo que consiste en:

30 (i) polipéptidos según la reivindicación 2;

(ii) polinucleótidos según la reivindicación 1; y

(iii) células presentadoras de antígeno que expresan un polinucleótido según la reivindicación 1;

35 en unas condiciones y durante un tiempo suficiente para permitir la estimulación y/o la expansión de células T;

(b) detectar en la muestra la magnitud de dicha estimulación y/o expansión de células T; y

40 (c) comparar la magnitud de dicha estimulación y/o expansión con un valor de corte predeterminado, y a partir de ello, identificar antígenos bacterianos asociados con una enfermedad inflamatoria del intestino en el mamífero.

45 25. Procedimiento según la reivindicación 24, en el que dicho mamífero es un ratón, en el que opcionalmente dicho ratón es de una cepa seleccionada de entre el grupo que consiste en C3H/HeJ Bir, BALB/c IL-10  $-/-$ , B6 IL-10  $-/-$ , B10 IL-10  $-/-$ , MDR1a  $-/-$ , TCR $\alpha$   $-/-$ , IL-2  $-/-$ , IL-2R  $-/-$ , ratones con colitis inducida por DSS (dextrano-sulfato de sodio), Gaai  $-/-$  y ratones transgénicos CD45 RB, siendo preferentemente dicho ratón de la cepa C3H/HeJ Bir.

50 26. Procedimiento de monitorización de la progresión de una enfermedad inflamatoria del intestino en un mamífero, que comprende las etapas siguientes:

(a) poner en contacto una muestra biológica obtenida de dicho mamífero que comprende anticuerpos con un polipéptido según la reivindicación 2;

(b) detectar en la muestra una cantidad de anticuerpo que se une al polipéptido; y

55 (c) repetir las etapas (a) y (b) utilizando una muestra biológica obtenida del mamífero en un punto temporal posterior; y

60 (d) comparar la cantidad de anticuerpo unido en la etapa (b) con la cantidad de anticuerpo unido en la etapa (c) y a partir de ello, monitorizar la progresión de una enfermedad inflamatoria del intestino en el mamífero.

65 27. Procedimiento según la reivindicación 17 o la reivindicación 19, en el que dicha muestra biológica se selecciona de entre el grupo que consiste en sueros, heces, tejido u otro material obtenido mediante colonoscopia, tejido u otro material obtenido mediante ileoscopia, tejido u otro material obtenido mediante esofagogastroduodenoscopia (EGP), y tejido u otro material obtenido mediante cirugía.

28. Kit de diagnóstico, que comprende:

(a) al menos un anticuerpo según la reivindicación 10 y un reactivo de detección, en el que el reactivo de detección comprende un grupo indicador; o

5 (b) al menos uno o más polipéptidos según la reivindicación 2 y un reactivo de detección que comprende un grupo indicador y opcionalmente inmovilizado sobre un soporte sólido.

29. Kit según la reivindicación 28, en el que el reactivo de detección comprende un anticuerpo anti-inmunoglobulina, proteína G, proteína A o una lectina.

10 30. Kit según la reivindicación 28, en el que el grupo indicador se selecciona de entre el grupo que consiste en grupos radiactivos, grupos fluorescentes, grupos luminiscentes, enzimas, biotina y tintes.

15 31. Composición inmunogénica que comprende un polipéptido según la reivindicación 2, un antígeno y un portador adecuado.

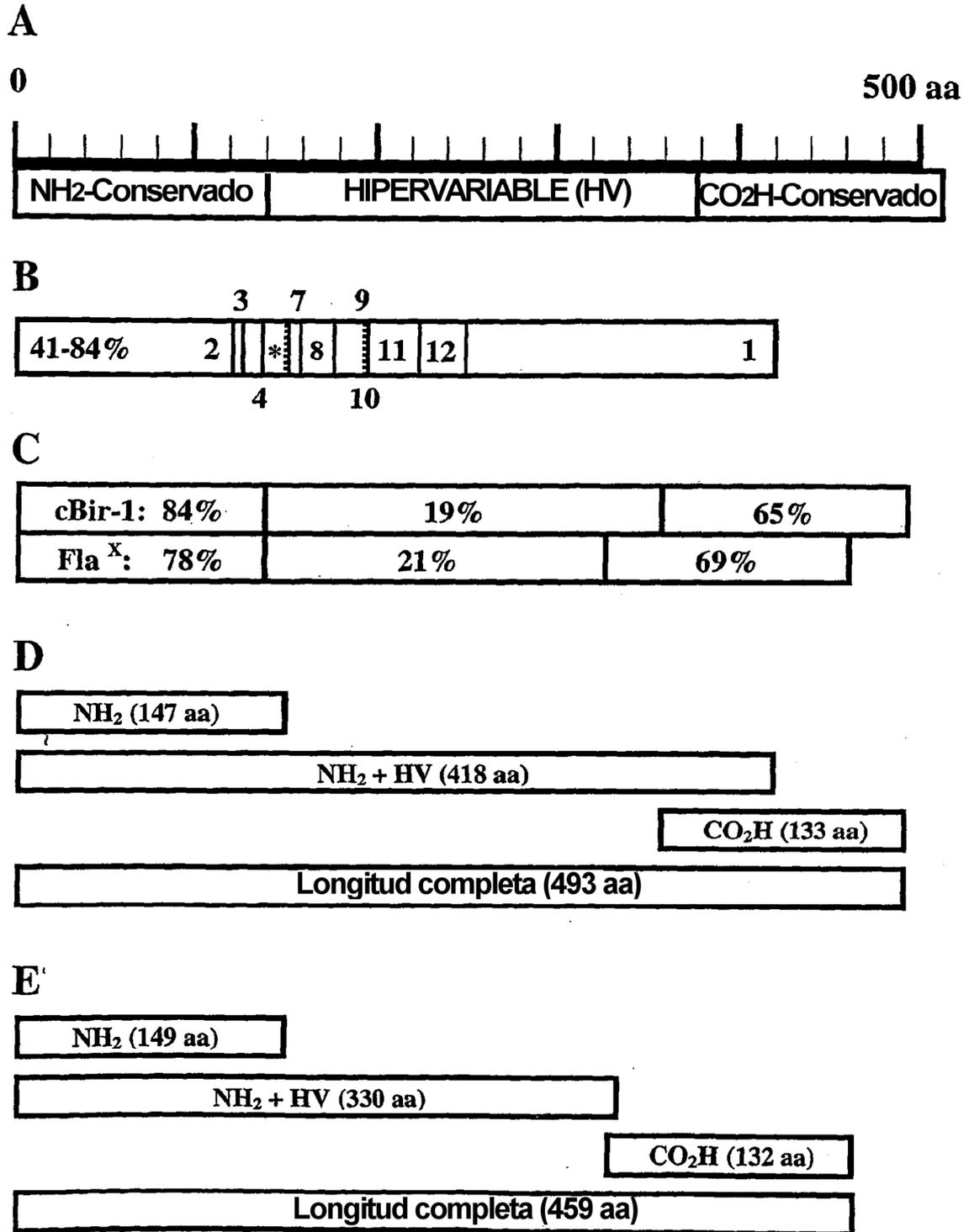


Figura 1

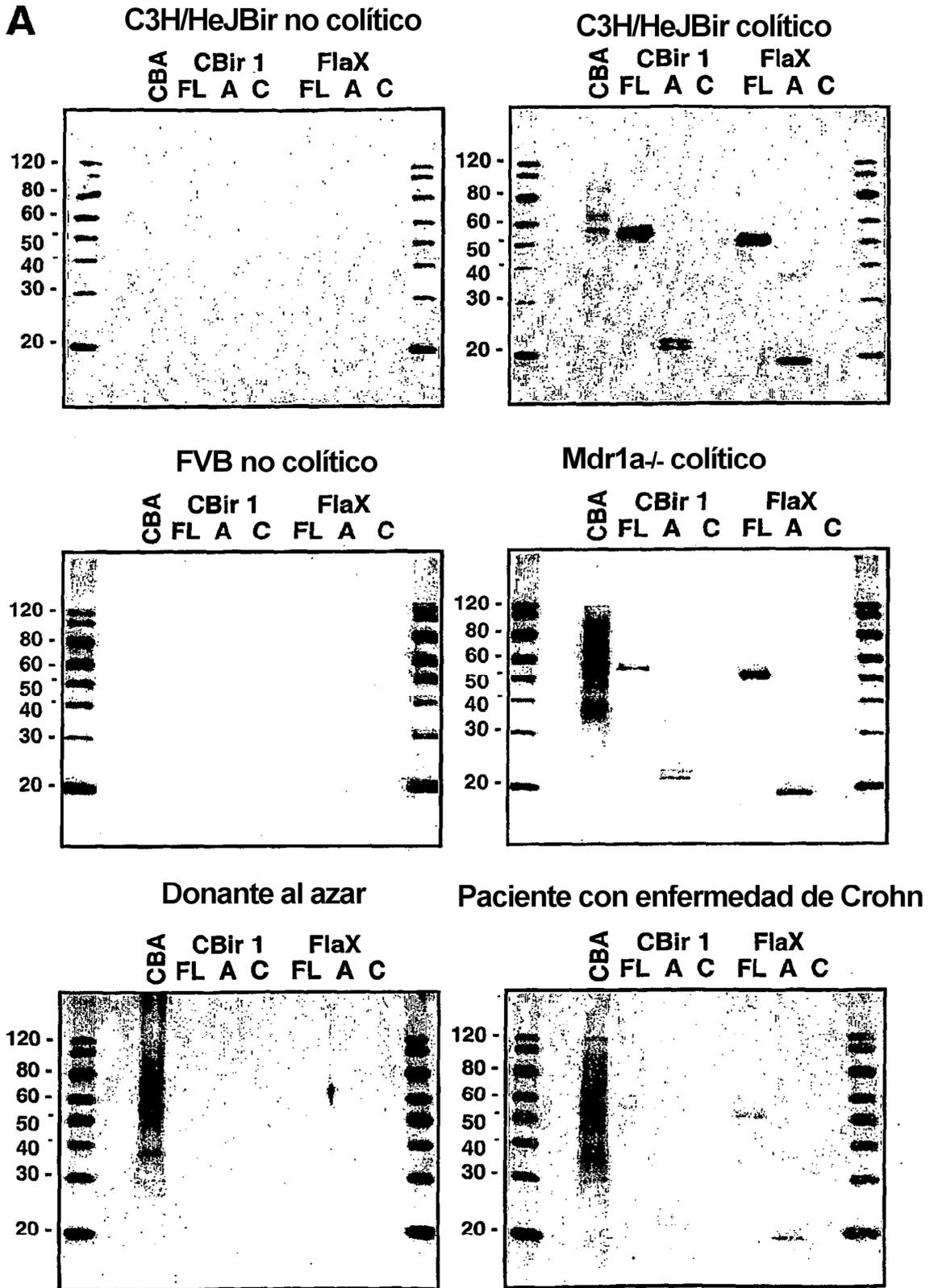
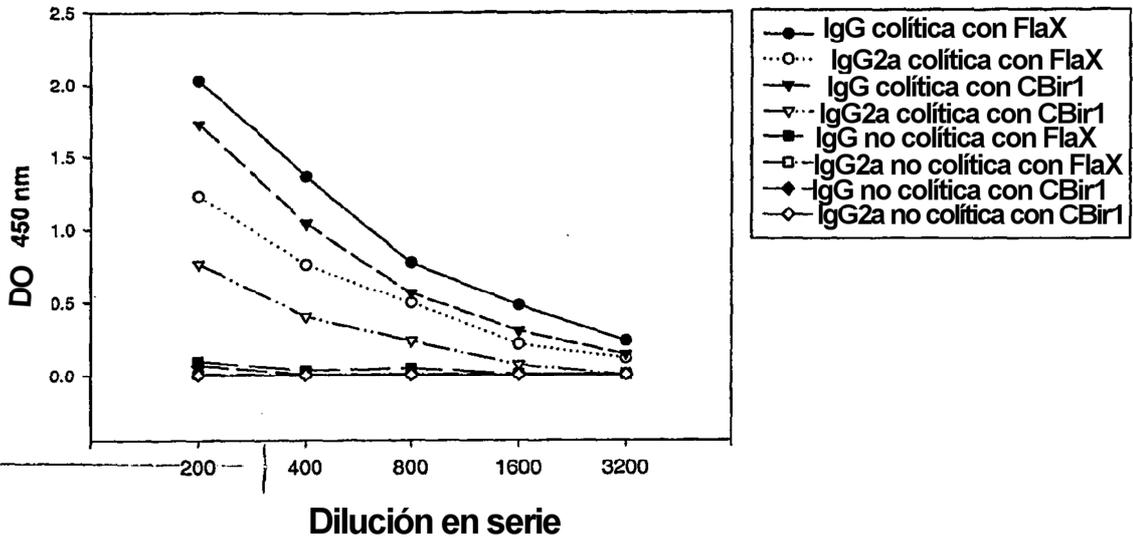


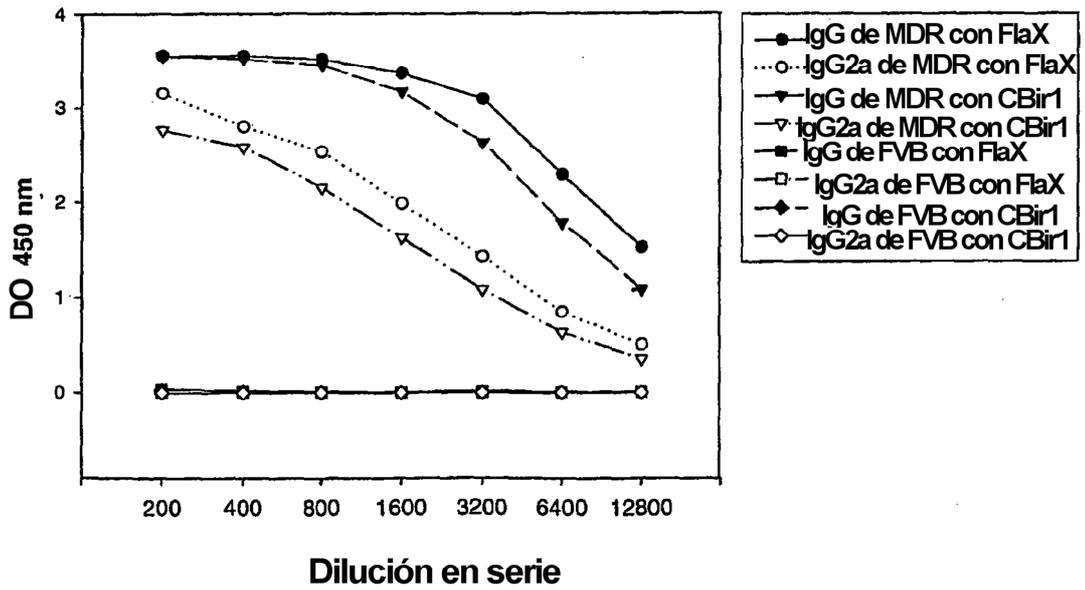
Figura 2A

**B**

**Suero de C3H/HeJBir**



**Combinaciones de suero de MDR frente a FVB**



**Figura 2B**

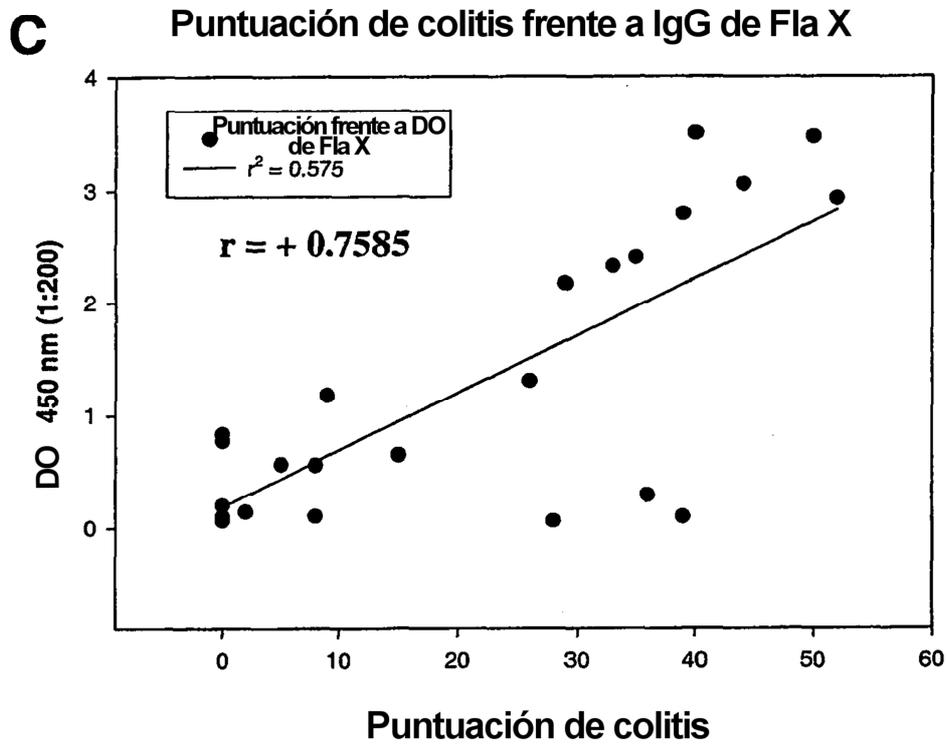
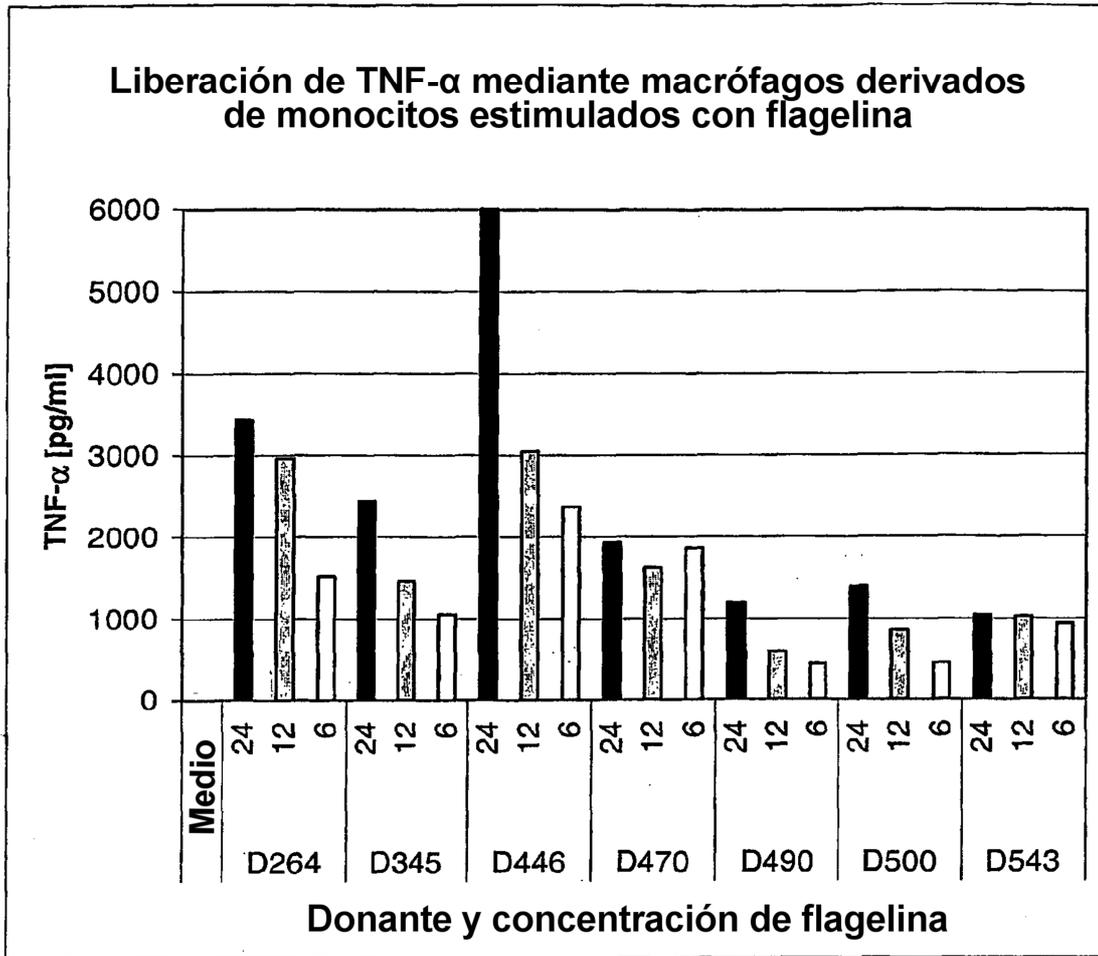
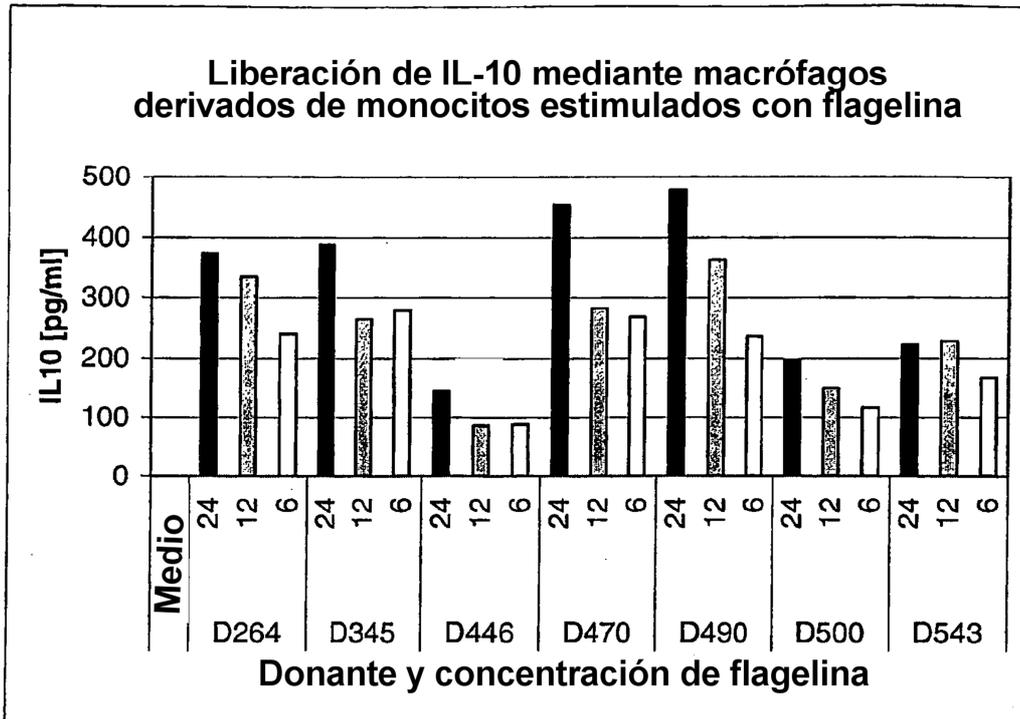


Figura 2c



**Figura 3A**



**Figura 3B**