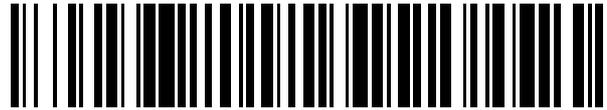


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 524 570**

51 Int. Cl.:

C07K 14/35 (2006.01)

A61K 39/04 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **27.04.2006 E 11190079 (1)**

97 Fecha y número de publicación de la concesión europea: **01.10.2014 EP 2426141**

54 Título: **Procedimiento de prevención o tratamiento de infección por M. tuberculosis**

30 Prioridad:

27.02.2006 US 777017 P

29.04.2005 US 676549 P

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

10.12.2014

73 Titular/es:

**GLAXOSMITHKLINE BIOLOGICALS S.A. (100.0%)
rue de l'Institut, 89
1330 Rixensart, BE**

72 Inventor/es:

MARCHAND, MARTINE

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 524 570 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

Procedimiento de prevención o tratamiento de infección por *M. tuberculosis*

Campo de la invención

La presente invención se refiere a proteínas de fusión M72 y aspectos relacionados.

5 Antecedentes de la invención

La tuberculosis es una enfermedad infecciosa crónica causada por la infección con *M. tuberculosis* y otras especies de *Mycobacterium*. Es una enfermedad principal en los países en desarrollo, así como un problema creciente en las áreas desarrolladas del mundo, con aproximadamente 8 millones de nuevos casos y 3 millones de muertes cada año. Aunque la infección puede ser asintomática durante un período de tiempo considerable, la enfermedad se manifiesta más comúnmente como una inflamación aguda de los pulmones, dando como resultado fiebre y una tos no productora. Si no se trata, resultan típicamente complicaciones graves y la muerte.

Aunque la tuberculosis se puede controlar en general empleando terapia prolongada con antibióticos, este tratamiento no es suficiente para evitar la propagación de la enfermedad. Los individuos infectados pueden ser asintomáticos, pero contagiosos, durante algún tiempo. Además, aunque es crítico el cumplimiento con el régimen de tratamiento, es difícil monitorear el comportamiento del paciente. Algunos pacientes no terminan el periodo de tratamiento, lo que puede conducir a un tratamiento inefectivo y al desarrollo de resistencia a los fármacos. Inclusive cuando se ha terminado un curso completo de tratamiento, la infección con *M. tuberculosis* no se erradica del individuo infectado, sino que permanece como una infección latente que se puede reactivar.

Con el objeto de controlar la propagación de la tuberculosis, es de primordial importancia una vacunación efectiva y un diagnóstico precoz preciso de la enfermedad. En la actualidad, la vacunación con bacterias vivas es el procedimiento más eficiente para inducir una inmunidad protectora. La micobacteria más común empleada para este propósito es *Bacillus Calmette-Guerin* (BCG), una cepa avirulenta de *M. bovis*. Sin embargo, la seguridad y eficacia de BCG es una fuente de controversia y algunos países, tales como los Estados Unidos, no vacunan al público en general con este agente.

El diagnóstico de la tuberculosis se logra comúnmente utilizando una prueba de piel, la que involucra la exposición intradérmica a la tuberculina PPD (derivado purificado de proteína). Las respuestas de las células-T específicas del antígeno dan como resultado una induración mensurable en el sitio de inyección a las 48-72 horas después de la inyección, lo que indica la exposición a los antígenos micobacterianos. Sin embargo, la sensibilidad y la especificidad han sido un problema con esta prueba y los individuos vacunados con BCG no se pueden distinguir de los individuos infectados.

Aunque se ha demostrado que los macrófagos actúan como los efectores principales de la inmunidad a *Mycobacterium*, las células-T son los inductores predominantes de esta inmunidad. El papel esencial de las células-T en la protección contra la infección por *Mycobacterium* se ilustra por la frecuente aparición de infección por *Mycobacterium* en los pacientes de SIDA, debido al agotamiento de las células-T CD4⁺ asociado con la infección por el virus de inmunodeficiencia humana (VIH). Se ha demostrado que las células-T CD4⁺ que reaccionan con *Mycobacterium* son potentes productores del interferón- γ (IFN- γ), que, a su vez, se ha demostrado que desencadena los efectos anti-micobacterianos de los macrófagos en los ratones. Aunque está menos claro el papel del IFN- γ en los seres humanos, los estudios han demostrado que la 1,25-dihidroxi-vitamina D3, ya sea sola o en combinación con IFN- γ o con el factor de necrosis tumoral- α , activa los macrófagos humanos para inhibir la infección por *M. tuberculosis*. Adicionalmente, se sabe que el IFN- γ estimula los macrófagos humanos para hacer la 1,25-dihidroxi-vitamina D3. De una manera similar, se ha demostrado que la interleucina-12 (IL-12) tiene un papel en el estímulo de la resistencia a la infección por *M. tuberculosis*. Para una revisión de la inmunología de la infección por *M. tuberculosis*, véase Chan y Kaufmann, *Tuberculosis: Pathogenesis, Protection and Control* (Bloom, editor, 1994), *Tuberculosis* (2ª edición, Rom y Garay, editores, 2003) y Harrison's Principles of Internal Medicine, Capítulo 150, páginas 953-966 (16ª edición, Braunwald y colaboradores, editores, 2005).

Los documentos WO03/070187 y WO01/98460 describen proteínas de fusión de *Mycobacterium tuberculosis*.

El documento WO03/072792 describe construcciones de proteínas de fusión heterólogas que comprenden un antígeno de leishmania.

Descripción de las secuencias enumeradas

- 50 SEC ID N.º:1: Mtb72f con etiqueta 6 His N-terminal (ADN).
 SEC ID N.º:2: Mtb72f con etiqueta 6 His N-terminal (proteína).

SEC ID N.º:3: M72 (variante de Mtb72f) con inserción de 2 His N-terminal (ADN).

SEC ID N.º:4: M72 (variante de Mtb72f) con inserción de 2 His N-terminal (proteína).

SEC ID N.º:5: Mtb72f sin inserción de His N-terminal (ADN).

SEC ID N.º:6: Mtb72f sin inserción de His N-terminal (proteína).

5 Breve descripción de la invención

La presente invención proporciona un polipéptido que comprende la secuencia de aminoácidos de SEC ID N.º: 4, un polinucleótido que comprende una secuencia de ácidos nucleicos que codifican la secuencia de aminoácidos de SEC ID N.º: 4 y composiciones farmacéuticas relacionadas.

Breve descripción de los dibujos

10 La Figura 1 muestra una representación gráfica del modelo de reactivación de *M. tuberculosis* en ratones Swiss Webster (SWR/J). La figura muestra los puntos del tiempo para la infección, el tratamiento con quimioterapia (50 miligramos de rifampina/85 miligramos de isoniazida por litro del agua para beber), las inmunizaciones y la enumeración de la carga bacteriana/las unidades formadoras de colonias (CFU).

15 La Figura 2 muestra las respuestas inmunes de los anticuerpos IgG1 e IgG2a en ratones SWR/J infectados con *M. tuberculosis* tratados con quimioterapia y luego inmunizados con Mtb72f. Los ratones se dejaron sin tratamiento, se trataron con quimioterapia (50 miligramos de rifampina/85 miligramos de isoniazida por litro del agua potable), o bien se trataron con quimioterapia y se inmunizaron tres veces intramuscularmente con 8 microgramos por dosis de Mtb72f formulada sin adyuvante. 10 días después de la última inmunización, los ratones se sangraron y se probaron los sueros para determinar la respuesta del anticuerpo anti-Mtb72f para
20 ambos isótopos IgG1 (rojo) e IgG2a (negro) mediante ELISA.

La Figura 3 muestra las respuestas inmunes de los anticuerpos IgG1 e IgG2a en los ratones SWR/J infectados por *M. tuberculosis* tratados con quimioterapia y luego inmunizados con Mtb72f. Los ratones se dejaron sin tratamiento, se trataron con quimioterapia (50 miligramos de rifampina/85 miligramos de isoniazida por litro del agua potable), o se trataron con quimioterapia y se inmunizaron tres veces intramuscularmente con 8
25 microgramos por dosis de Mtb72f formulada con el adyuvante AS01B. 10 días después de la última inmunización, los ratones se sangraron y se probaron los sueros para determinar la respuesta del anticuerpo anti-Mtb72f para ambos isótopos IgG1 (rojo) e IgG2a (negro) mediante ELISA.

La Figura 4 muestra las respuestas del interferón-gamma (IFN- γ) en los ratones SWR/J infectados por *M. tuberculosis* tratados con quimioterapia y luego inmunizados con Mtb72f. Se obtuvieron células de bazo de los
30 ratones en diferentes puntos del tiempo y se estimularon *in vitro* durante 3 días con 10 microgramos/mililitro ya sea de rMtb72f o bien de los componentes (Mtb32_C y Mtb39), como se indica. Como controles, también se estimularon cultivos de esplenocitos ya sea con PPD (3 microgramos/mililitro), lisado de BCG (10 microgramos/mililitro), conA (3 microgramos/mililitro), o el medio solo. Subsiguientemente se midió la producción de IFN- γ mediante ELISA.

35 La Figura 5 muestra las respuestas de IFN- γ en los ratones SWR/J infectados por *M. tuberculosis* tratados con quimioterapia y luego inmunizados con Mtb72f. Se obtuvieron células de bazo de los ratones en diferentes puntos del tiempo y se estimularon *in vitro* durante 3 días con 10 microgramos/mililitro ya sea de rMtb72f, o bien de los componentes (Mtb32_C y Mtb39), como se indica. Como los controles, también se estimularon cultivos de esplenocitos ya sea con PPD (3 microgramos/mililitro), lisado de BCG (110 microgramos/mililitro), conA (3
40 microgramos/mililitro), o el medio solo. Subsiguientemente se midió la producción de IFN- γ mediante ELISA.

La Figura 6 muestra las respuestas de células-T CD4⁺ y de citocinas de IFN- γ en los ratones SWR/J infectados por *M. tuberculosis* tratados con quimioterapia y luego inmunizados con Mtb72f. Se obtuvieron células de bazo de los ratones en diferentes puntos del tiempo y se estimularon *in vitro* durante la noche con 10 microgramos/mililitro de rMtb72f. Luego las células se tiñeron para CD4 e IFN- γ . Como un control, también se estimularon cultivos de esplenocitos con el medio solo. Subsiguientemente se midió la producción de IFN- γ + específica de células-T CD4⁺, mediante tinte de citocina intracelular (ICS).
45

La Figura 7 muestra un resumen tabular de los valores de la producción de IFN- γ + específico de células-T CD4+ y CD8+ en el Día 120 después de la infección por Mtb. Se obtuvieron células de bazo de grupos de ratones dejados sin tratamiento, tratados durante 30, 60, o 90 días con quimioterapia de combinación, o tratados con quimioterapia de combinación como un auxiliar a la vacuna de Mtb72f. Los esplenocitos se estimularon se estimularon *in vitro* durante la noche con 10 microgramos/mililitro de rMtb72f. Luego las células se tiñeron para determinar CD4, CD8, o IFN- γ . Como un control, los cultivos de esplenocitos también se estimularon con el medio solo. Subsiguientemente se midió la producción de IFN- γ + específico de células-T CD4+ y CD8+, mediante tinte
50

con citocina intracelular.

5 La Figura 8 muestra la supervivencia de los ratones SWR/J infectados por *M. tuberculosis* tratados con quimioterapia y luego inmunizados con Mtb72f. Los ratones se infectaron mediante un aerosol con 50 a 100 unidades formadoras de colonias de MtbH37Rv y se inició la quimioterapia (50 miligramos de rifampina/85 miligramos de isoniazida por litro del agua para beber) en un subconjunto de ratones 30 días después. La quimioterapia se continuó durante 60 días.

La mitad de los ratones que recibieron quimioterapia se inmunizaron tres veces intramuscularmente con 8 microgramos por dosis de Mtb72f formulada con el adyuvante AS01B.

10 La Figura 9 muestra la supervivencia de los ratones SWR/J infectados por *M. tuberculosis* tratados con quimioterapia y luego inmunizados con Mtb72f. Los ratones se infectaron mediante un aerosol con 50 a 100 unidades formadoras de colonias de MtbH37Rv y se inició la quimioterapia (50 miligramos de rifampina/85 miligramos de isoniazida por litro del agua para beber) en un subconjunto de ratones 30 días después. La quimioterapia se continuó durante 30, 60, o 90 días en subconjuntos separados de ratones. La mitad de los ratones que recibieron quimioterapia se inmunizaron tres veces intramuscularmente con 8 microgramos por dosis de Mtb72f formulada con el adyuvante AS01B.

Descripción detallada de las realizaciones específicas

La presente invención proporciona un polipéptido que comprende la secuencia de aminoácidos de SEC ID N.º: 4, un polinucleótido que comprende una secuencia de ácidos nucleicos que codifican la secuencia de aminoácidos de SEC ID N.º: 4 y composiciones farmacéuticas relacionadas.

20 Los ácidos nucleicos y los polipéptidos de fusión de la presente invención pueden comprender además otros componentes diseñados para mejorar su antigenicidad, o para mejorar estos antígenos en otros aspectos. Las composiciones farmacéuticas, polipéptidos y ácidos nucleicos de la invención, pueden comprender copias adicionales de antígenos, o polipéptidos heterólogos adicionales de *Mycobacterium sp.*, tales como el antígeno MTB8.4, el antígeno MTB9.8, el antígeno MTB9.9, el antígeno MTB40, el antígeno MTB41, el antígeno ESAT-6, el antígeno complejo MTB85, el antígeno α -cristalino, o el antígeno NS1. De una manera alternativa o adicional, las composiciones farmacéuticas, polipéptidos y ácidos nucleicos de la invención, pueden comprender copias adicionales de antígenos, o polipéptidos adicionales a partir de *Mycobacterium sp.*, tales como Ag85B o MTCC#2. Las composiciones farmacéuticas, polipéptidos y ácidos nucleicos de la invención también pueden comprender polipéptidos adicionales de otras fuentes. Por ejemplo, las composiciones farmacéuticas y proteínas de fusión de la invención pueden incluir polipéptidos o ácidos nucleicos que codifiquen polipéptidos, en donde el polipéptido mejore la expresión del antígeno, por ejemplo, NS1, una proteína de virus de influenza (véanse, por ejemplo, las publicaciones internacionales WO99/40188 y WO93/04175). Los ácidos nucleicos de la invención se pueden diseñar basándose en la preferencia de codones en una especie de elección, por ejemplo en seres humanos.

35 Las composiciones de proteínas de fusión normalmente comprenden uno o más adyuvantes, por ejemplo AS01B (monofosforil-lípido A (MPL) y QS21 en una formulación de liposoma; véase la Publicación de Patente de los EE.UU. n.º: 2003/0143240); AS02A (3D-MPL y QS21 y una emulsión de aceite en agua; véase Bojang y colaboradores, *Lancet* (2001) 358:1927); ENHANZYN (Detox); 3D-MPL; saponinas incluyendo Quil A y sus componentes por ejemplo QS21 y miméticos de saponina; CWS; TDM; AGP; oligonucleótidos inmunoestimulantes, por ejemplo CPG; Leif; y derivados de los mismos. En una realización preferida, se administra un polipéptido de fusión con uno o más adyuvantes seleccionados a partir del grupo que consiste en 3D-MPL y QS21 en una formulación de liposoma por ejemplo AS01B y MPL y QS21 y una emulsión de agua en aceite (por ejemplo AS02A). Los adyuvantes AS01B y AS02A se describen adicionalmente en Pichyangkul y colaboradores, *Vaccine* (2004) 22:3831-40.

45 Cuando se suministra el antígeno como un ácido nucleico, se puede administrar, por ejemplo, en un vector viral (es decir, un vector de adenovirus), o en una célula huésped de bacteria mutante (es decir, una célula huésped avirulenta mutante de *Mycobacterium*, *Lactobacillus*, o *Bacillus*, incluyendo *Bacillus Calmette-Guerin* (BCG) y *Lactococcus lactis*).

Las composiciones de Mtb72f normalmente se administran a seres humanos, pero son efectivas en otros mamíferos, incluyendo mamíferos domésticos (es decir, perros, gatos, conejos, ratas, ratones, conejillos de indias, hámsteres, chinchillas) y mamíferos agrícolas (es decir, reses, cerdos, ovejas, cabras, caballos).

50 En la nomenclatura de la solicitud, Ra35 se refiere al extremo N-terminal de Mtb32A (Ra35FL), la porción de la SEC ID N.º: 2 dada a conocer en la presente solicitud, correspondiente a los residuos 535-729.

Ra12 se refiere al extremo C-terminal de Mtb32A (Ra35FL), la porción de la SEC ID N.º:2 que se da a conocer en la presente solicitud, correspondiente a los residuos 8-139.

Mtb39 (TbH9) se refiere a una secuencia esencialmente como la que se da a conocer como la SEC ID N.º:106

(ADNc de longitud completa) y como la SEC ID N.º:107 (proteína de longitud completa) en las solicitudes de patente de los EE.UU. n.º: 08/658.800, n.º: 08/659.683, n.º: 08/818.112 y n.º: 08/818.111 y en las solicitudes internacionales WO97/09428 y WO97/09429. TbH9 se refiere a la porción de la SEC ID N.º:2 que se da a conocer en la presente solicitud, correspondiente a los residuos 143-532.

5 Ra12-TbH9-Ra35 (Mtb72f), cuya secuencia se da a conocer como la SEC ID N.º:1 o la SEC ID N.º:5 (ADN) y la SEC ID N.º: 2 o la SEC ID N.º:6 (proteína) en la presente solicitud, así como en la solicitud de patente de los EE.UU. n.º: 09/223.040 y en el documento TCP PCT/US99/07717. Las secuencias de la SEC ID N.º:1 y de la SEC ID N.º: 2 incluyen la marca His de los seis residuos His.

10 M72 es un mutante de Mtb72f en donde el residuo de serina en el aminoácido correspondiente a la posición 710 de la SEC ID N.º:2 ha sido cambiado hasta Ala (así como cuatro residuos His que se han removido de la marca His en el extremo N-terminal), cuya secuencia se da a conocer como la SEC ID N.º:3 (ADN) y la SEC ID N.º:4 (proteína) en la presente solicitud. Una variante de estas secuencias, en donde la proteína tiene una marca His de los seis residuos His, se da a conocer en la solicitud de patente de los EE.UU. n.º: 09/597.796 y en la solicitud internacional del TCP PCT/USO1/19959. En virtud del reemplazo de Ser 710 con Ala, se cree que M72 es más resistente a la autólisis que Mtb72f.

15 Lo siguiente proporciona secuencias de algunos antígenos individuales utilizados en las composiciones y proteínas de fusión de la invención:

20 Mtb8.4 (DPV), cuya secuencia se da a conocer como la SEC ID N.º:101 (ADNc) y la SEC ID N.º:102 (proteína) en las solicitudes de patente de los EE.UU. n.º: 08/658.800, n.º: 08/659.683, n.º: 08/818.112 y n.º: 08/818.111 y en las solicitudes internacionales WO97/09428 y WO97/09429.

Mtb9.8 (MSL), cuya secuencia se da a conocer como la SEC ID N.º:12 (ADN), la SEC ID N.º:109 (secuencia de aminoácidos predicha) y las SEC ID N.ºs:110 a 124 (péptidos) en las solicitudes de patente de los EE.UU. n.º: 08/859.381, n.º: 08/858.998, n.º: 09/073.009 y n.º: 09/073.010 y en las solicitudes internacionales del TCP PCT/US98/10407 y PCT/US98/10514.

25 Mtb9.9A (MTI, también conocido como MTI-A), cuya secuencia se da a conocer como la SEC ID N.º:3 y la SEC ID N.º:4 (ADN) y como la SEC ID N.º:29 y las SEC ID N.ºs:51 a 66 (péptido del marco de lectura abierta para MTI) en las solicitudes de patente de los EE.UU. n.º: 08/859.381, n.º: 08/858.998, n.º: 09/073.009 y n.º: 09/073.010 y en las solicitudes internacionales del TCP PCT/US98/10407 y PCT/US98/10514. También existen otras dos variantes de MTI, denominadas como MTI-B y MTI-C.

30 Mtb40 (HTCC#1), cuya secuencia se da a conocer como la SEC ID N.º:137 (ADNc) y la SEC ID N.º:138 (secuencia de aminoácidos predicha), en las solicitudes de patente de los EE.UU. n.º: 09/073.009 y n.º: 09/073.010 y en las solicitudes internacionales del TCP PCT/US98/10407 y PCT/US98/10514.

35 Mtb41 (MTCC#2), cuya secuencia se da a conocer como la SEC ID N.º:140 (ADNc) y la SEC ID N.º:142 (secuencia de aminoácidos predicha), en las solicitudes de patente de los EE.UU. n.º: 09/073.009 y n.º: 09/073010 y en las solicitudes internacionales del TCP PCT/US98/ 10407 y PCT/US98/10514.

ESAT-6, cuya secuencia se da a conocer como la SEC ID N.º:103 (ADN) y la SEC ID N.º:104 (secuencia de aminoácidos predicha), en la solicitud de patente de los EE.UU. n.º: 09/072.967. La secuencia de ESAT-6 también se da a conocer en la Patente de los EE.UU. n.º: 5.955.077.

40 El antígeno α -cristalino, cuya secuencia se da a conocer en Verbon y colaboradores, *J. Bad.* 174:1352- 1359 (1992).

El antígeno del complejo 85, cuya secuencia se da a conocer en Content y colaboradores, *Infect. & Immunol.* 59:3205-3212 (1991).

Cada una de las secuencias anteriores también se da a conocer en Cole y colaboradores, *Nature* 393:537 (1998) y se pueden encontrar, por ejemplo, en <http://www.sanger.ac.uk> y <http://www.pasteur.fr/mycdb/>.

45 Las secuencias anteriores se dan a conocer en las solicitudes de patente de los EE.UU. n.ºs: 08/523.435; 08/523.436; 08/658.800; 08/659.683; 08/818.111; 08/818.112; 08/942.341; 08/942.578; 08/858.998; 08/859.381; 09/056.556; 09/072-596; 09/072.967; 09/073.009; 09/073.010; 09/223.040; 09/287.849; y en las solicitudes de patente del TCP PCT/US98/10407, PCT/US98/10514, PCT/US99/03265, PCT/US99/03268, PCT/US99/07717 y en las solicitudes internacionales WO97/09428 y WO97/09429, WO98/16645, WO98/16646.

50 Los antígenos descritos en la presente memoria incluyen variantes polimórficas y variaciones conservadoramente modificadas, así como homólogos de *Mycobacterium* inter-cepas e inter-especies. Además, los antígenos descritos en la presente memoria incluyen subsecuencias o secuencias truncadas. Las proteínas de fusión también pueden

contener polipéptidos adicionales, opcionalmente péptidos heterólogos a partir de *Mycobacterium* o de otras fuentes. Estos antígenos se pueden modificar, por ejemplo, mediante la adición de secuencias peptídicas enlazadoras, como se describen más adelante. Los péptidos enlazadores se pueden insertar entre uno o más componentes que forman cada una de las proteínas de fusión.

5 Definiciones

El término “reactivación de tuberculosis” se refiere a la manifestación posterior de los síntomas de la enfermedad en un individuo que se pruebe como positivo en una prueba de tuberculina, pero que no tenga los síntomas aparentes de la enfermedad. El individuo está infectado con *M. tuberculosis* y puede o no haber manifestado previamente los síntomas de la enfermedad activa que se hubieran tratado suficientemente para llevar a la tuberculosis hasta un estado inactivo o latente. Sin embargo, se pueden iniciar procedimientos para la prevención o el tratamiento de la reactivación de tuberculosis en un individuo que manifieste síntomas activos de la enfermedad.

“Tuberculosis primaria” se refiere a la enfermedad clínica (manifestación de los síntomas de la enfermedad) directamente a continuación de la infección con *M. tuberculosis*. Véase *Harrison's Principles of Internal Medicine*, Capítulo 150, páginas 953-966 (16ª Edición, Braunwald y colaboradores, editores, 2005).

15 “Tuberculosis secundaria” o “tuberculosis post-primaria”, se refiere a la reactivación de una infección por *M. tuberculosis* durmiente, inactiva, o latente. Véase *Harrison's Principles of Internal Medicine*, *supra*.

Una “infección activa de *M. tuberculosis*” se refiere a una infección por *M. tuberculosis* con los síntomas de la enfermedad manifestados. Párrafo [0045]

20 Una “infección inactiva, durmiente, o latente de *M. tuberculosis*” se refiere a una infección por *M. tuberculosis* sin los síntomas manifestados de la enfermedad.

Una infección por *M. tuberculosis* “resistente a fármacos” se refiere a una infección por *M. tuberculosis* en donde la cepa infecciosa no se mantiene estática ni se elimina mediante (es resistente a) uno o más de los agentes quimioterapéuticos denominados de la “línea frontal” efectivos en el tratamiento de una infección por *M. tuberculosis* (por ejemplo, isoniazida, rifampina, etambutol, estreptomycin y pirazinamida).

25 Una infección por *M. tuberculosis* “resistente a múltiples fármacos” se refiere a una infección por *M. tuberculosis* en donde la cepa infecciosa es resistente a dos o más agentes quimioterapéuticos de la “línea frontal” efectivos para el tratamiento de una infección por *M. tuberculosis*.

Un “agente quimioterapéutico efectivo en el tratamiento de una infección por *M. tuberculosis*” se refiere a los agentes farmacológicos conocidos y utilizados en la técnica para el tratamiento de infecciones por *M. tuberculosis*. Los agentes farmacológicos ejemplificados utilizados para tratar infecciones por *M. tuberculosis* incluyen, pero no se limitan a, amicacina, ácido amino-salicílico, capreomicina, cicloserina, etambutol, etionamida, canamicina, pirazinamida, rifamicinas (es decir, rifampina, rifapentina y rifabutina), estreptomycin, ofloxacina, ciprofloxacina, claritromicina, azitromicina y fluoro-quinolonas. Los agentes quimioterapéuticos de “primera línea” utilizados para el tratamiento de una infección por *M. tuberculosis* que no sea resistente a fármacos incluyen isoniazida, rifampina, etambutol, estreptomycin y pirazinamida. Los agentes quimioterapéuticos de “segunda línea” utilizados para el tratamiento de una infección por *M. tuberculosis* que haya demostrado resistencia a uno o más fármacos de la “primera línea” incluyen ofloxacina, ciprofloxacina, etionamida, ácido amino-salicílico, cicloserina, amicacina, canamicina y capreomicina. Estos agentes farmacológicos se revisan en el Capítulo 48 de *Goodman and Gilman's The Pharmacological Basis of Therapeutics*, Hardman y Limbird, editores, 2001.

40 “FL” se refiere a longitud completa, es decir, un polipéptido que tiene la misma longitud que la del polipéptido de tipo silvestre.

“Marca His” se refiere a un cordón de residuos His, típicamente seis residuos que se insertan en el extremo N-terminal, en general inmediatamente después del residuo Met de inicio, o de otra manera en el termino C. Normalmente son heterólogos para la secuencia nativa, pero se incorporan debido a que facilitan el aislamiento al mejorar el enlace de la proteína con las resinas de cromatografía de afinidad con metal inmovilizado (IMAC). Hablando en términos generales, la presencia o la ausencia de una marca His no es significativa desde el punto de vista de provocar una respuesta inmune útil contra la proteína antigénica que se vaya a provocar. En el caso de que se provoque una reacción inmune adversa contra la marca His misma, se considera mejor minimizar la longitud de la marca His, por ejemplo hasta 4 o menos residuos, en particular hasta 2 residuos.

50 El término “fragmento inmunógeno del mismo” se refiere a un polipéptido que comprende un epítipo que es reconocido por los linfocitos-T citotóxicos, los linfocitos-T auxiliares, o las células-B. Típicamente, un fragmento inmunógeno de Mtb72f será un polipéptido que contenga 500 o más aminoácidos, por ejemplo 600 o más aminoácidos, por ejemplo 700 o más aminoácidos. La invención también abarca una pluralidad de fragmentos, por ejemplo, fragmentos traslapados, que cubren juntos toda o sustancialmente toda (por ejemplo, 500 o más

aminoácidos, por ejemplo 600 o más aminoácidos, por ejemplo 700 o más aminoácidos) la secuencia de una proteína de fusión Mtb72f.

El término “especie de *Mycobacterium* del complejo de tuberculosis” incluye las especies que se considera tradicionalmente que provocan la enfermedad de tuberculosis, así como las especies ambientales y oportunistas de *Mycobacterium* que provocan tuberculosis y enfermedad pulmonar en los pacientes inmunocomprometidos, tales como los pacientes con SIDA, por ejemplo *M. tuberculosis*, *M. bovis*, o *M. africanum*, *BCG*, *M. avium*, *M. intracellulare*, *M. celatum*, *M. genavense*, *M. haemophilum*, *M. kansasii*, *M. simiae*, *M. vaccae*, *M. fortuitum* y *M. scrofulaceum* (véase, por ejemplo, *Harrison’s Principles of Internal Medicine*, Capítulo 150, páginas 953-966 (16ª Edición, Braunwald y colaboradores, editores, 2005). Párrafo [0052]

Un adyuvante se refiere a los componentes de una composición de vacuna o terapéutica, que aumentan la respuesta inmune específica al antígeno (véase, por ejemplo, Edelman, *AIDS Res. Hum Retroviruses* 8:1409-1411 (1992)). Los adyuvantes inducen respuestas inmunes del tipo Th1 y del tipo Th2. Las citocinas del tipo Th1 (por ejemplo, IFN- γ , IL-2, e IL-12) tienden a favorecer la inducción de la respuesta inmune mediada por células a un antígeno administrado, mientras que las citocinas del tipo Th2 (por ejemplo, IL-4, IL-5, IL-6, IL-10 y TNF- β) tienden a favorecer la inducción de las respuestas inmunes humorales. Los adyuvantes capaces de tener un estímulo preferencial de una respuesta inmune mediada por células Th1 se describen en las publicaciones internacionales WO 94/00153 Y WO 95/17209.

“Ácido nucleico” se refiere a los desoxirribonucleótidos o a los ribonucleótidos y polímeros de los mismos, ya sea en la forma de una sola cadena o de doble cadena. El término ácido nucleico se utiliza de una manera intercambiable con genes, ADNc, ARNm, oligonucleótido y polinucleótido.

Los términos “polipéptido”, “péptido” y “proteína”, se utilizan de una manera intercambiable en la presente memoria para referirse a un polímero de residuos de aminoácidos.

El término “aminoácido” se refiere a los aminoácidos que se presentan naturalmente y sintéticos, así como a los análogos de aminoácidos y a los miméticos de aminoácidos que funcionan de una manera similar a los aminoácidos que se presentan naturalmente. Los aminoácidos que se presentan naturalmente son aquellos codificados por el código genético, así como los aminoácidos que se modifican posteriormente, por ejemplo, hidroxil-prolina, γ -carboxi-glutamato y O-fosfoserina. Los análogos de aminoácidos se refieren a los compuestos que tienen la misma estructura química básica que el aminoácido que se presenta naturalmente, es decir, un átomo de carbono α que está enlazado con un hidrógeno, un grupo carboxilo, un grupo amino y un grupo R, por ejemplo homoserina, norleucina, sulfóxido de metionina, metionina-metil-sulfonio. Estos análogos tienen grupos R modificados (por ejemplo, norleucina) o estructuras base peptídicas modificadas, pero retienen la misma estructura química básica que un aminoácido que se presenta naturalmente. Los miméticos de aminoácidos se refieren a los compuestos químicos que tienen una estructura que es diferente de la estructura química general de un aminoácido, pero que funcionan de una manera similar al aminoácido que se presenta naturalmente.

Los aminoácidos pueden ser referidos en la presente memoria mediante sus símbolos de tres letras comúnmente conocidos, o mediante los símbolos de una letra recomendados por la Comisión de Nomenclatura Bioquímica IUPAC-IUB. De la misma manera, los nucleótidos pueden ser referidos por sus códigos de una sola letra comúnmente aceptados.

“Variantes conservadoramente modificadas” se aplica a las secuencias tanto de aminoácidos como de ácidos nucleicos. Con respecto a las secuencias de ácidos nucleicos particulares, las variantes conservadoramente modificadas se refieren a los ácidos nucleicos que codifican secuencias de aminoácidos idénticas o esencialmente idénticas, o en donde el ácido nucleico no codifica una secuencia de aminoácidos, hasta secuencias esencialmente idénticas. Debido a la degeneración del código genético, un gran número de ácidos nucleicos funcionalmente idénticos codifican cualquier proteína dada. Por ejemplo, los codones GCA, GCC, GCG y GCU codifican todos al aminoácido alanina. Por consiguiente, en cada posición en donde una alanina sea especificada por un codón, el codón se puede alterar hasta cualquiera de los codones correspondientes descritos, sin alterar el polipéptido codificado. Estas variaciones de ácidos nucleicos son “variaciones silenciosas”, que son una especie de variaciones conservadoramente modificadas. Cada secuencia de ácido nucleico de la presente memoria que codifique un polipéptido, también describe toda posible variación silenciosa del ácido nucleico. Un experto reconocerá que cada codón en un ácido nucleico (excepto AUG, que ordinariamente es el único codón para metionina y TGG, que ordinariamente es el único codón para triptófano) se puede modificar para proporcionar una molécula funcionalmente idéntica. De conformidad con lo anterior, cada variación silenciosa de un ácido nucleico que codifique un polipéptido es implícita en cada secuencia descrita.

Con respecto a las secuencias de aminoácidos, un experto reconocerá que las sustituciones, supresiones, o adiciones individuales a una secuencia de ácido nucleico, péptido, polipéptido, o proteína, que alteren, agreguen, o supriman un solo aminoácido o un pequeño porcentaje de aminoácidos en la secuencia codificada, es una “variante conservadoramente modificada”, en donde la alteración dé como resultado la sustitución de un aminoácido con un

aminoácido químicamente similar. Las tablas de sustitución conservadora que proporcionan aminoácidos funcionalmente similares son bien conocidas en este campo. Estas variantes conservadoramente modificadas son, además de y sin excluir, variantes polimórficas, homólogos inter-especies y alelos de la invención.

Los siguientes ocho grupos contienen cada uno aminoácidos que son sustituciones conservadoras unos por otros:

- 5 1) Alanina (A), glicina (G);
- 2) Ácido aspártico (D), ácido glutámico (E);
- 3) Asparagina (N), Glutamina (Q);
- 4) Arginina (R), Lisina (K);
- 5) Isoleucina (I), Leucina (L), Metionina (M), Valina (V);
- 10 6) Fenilalanina (F), Tirosina (Y), Triptófano (W);
- 7) Serina (S), Treonina (T); y
- 8) Cisteína (C), Metionina (M);

(véase, por ejemplo, Creighton, *Proteins* (1984)).

15 El término “heterólogo”, cuando se utiliza con referencia a porciones de un ácido nucleico, indica que el ácido nucleico comprende dos o más subsecuencias que no se encuentran en la misma relación unas con otras en la naturaleza. Por ejemplo, el ácido nucleico se produce típicamente de una manera recombinante, teniendo dos o más secuencias a partir de genes no relacionados, configuradas para formar un nuevo ácido nucleico funcional, por ejemplo, un promotor a partir de una fuente y una región codificante a partir de otra fuente. De una manera similar, una proteína heteróloga indica que la proteína comprende dos o más subsecuencias que no se encuentran en la misma relación unas con otras en la naturaleza (por ejemplo, una proteína de fusión).

20 “Polipéptido de fusión” o “proteína de fusión” se refiere a una proteína que tiene cuando menos dos polipéptidos de *Mycobacterium sp.* heterólogos covalentemente enlazados, ya sea directamente o bien por medio de un enlazador de aminoácidos. Los polipéptidos que forman la proteína de fusión típicamente se enlazan del extremo C-terminal al extremo N-terminal, aunque también se pueden enlazarse del extremo C-terminal al extremo N-terminal, del extremo N-terminal al extremo N-terminal, o del extremo N-terminal al extremo C-terminal. Los polipéptidos de la proteína de fusión pueden estar en cualquier orden. Este término también se refiere a las variantes conservadoramente modificadas, variantes polimórficas, alelos, mutantes, subsecuencias y homólogos inter-especies de los antígenos que forman la proteína de fusión. Los antígenos de *Mycobacterium tuberculosis* se describen en Cole y colaboradores, *Nature* 393:537 (1998), que da a conocer el genoma de *Mycobacterium tuberculosis* entero. La secuencia completa de *Mycobacterium tuberculosis* también se puede encontrar en <http://www.sanger.ac.uk> y en [http://www.pasteur.fr/mycdb/\(MycDB\)](http://www.pasteur.fr/mycdb/(MycDB)). Los antígenos de otras especies de *Mycobacterium* que corresponden a los antígenos de *M. tuberculosis* se pueden identificar, por ejemplo, utilizando algoritmos de comparación de secuencias, como se describen en la presente memoria, u otros procedimientos conocidos por los expertos en la materia, por ejemplo, ensayos de hibridación y ensayos de enlace de anticuerpos.

35 Las proteínas de fusión de ejemplo la presente invención son:

Las proteínas que comprenden o que consisten en la secuencia de la SEC ID N.º:4 (=M72);

y

Las proteínas de fusión que comprenden la secuencia de la SEC ID N.º:4 junto con uno o más antígenos de *M. tuberculosis* o un fragmento inmunógeno de cualquiera de las mismas.

40 Más específicamente la proteína de fusión es un polipéptido de SEC ID N.º:4.

Los ácidos nucleicos de la presente invención (por ejemplo, moléculas de ADN) son aquellos que codifican las proteínas de fusión anteriormente mencionadas. Las moléculas de ADN específicas que se pueden mencionar comprenden o consisten en la SEC ID N.º:1 o la SEC ID N.º:3.

45 El término “fusionado” se refiere a un enlace covalente entre dos polipéptidos en una proteína de fusión. Los polipéptidos típicamente se unen mediante un enlace peptídico, ya sea directamente unos con otros, o bien por medio de un enlazador de aminoácidos. Opcionalmente, los péptidos se pueden unir por medio de enlaces covalentes no peptídicos conocidos por los expertos en este campo.

La frase “se hibrida de una manera selectiva (o de una manera específica) a”, se refiere al enlace, duplexión, o

hibridación de una molécula solamente a una secuencia de nucleótidos particular según condiciones de hibridación restringentes, cuando esta secuencia está presente en una mezcla compleja (por ejemplo, ADN o ARN celular total o de biblioteca).

5 La frase "condiciones de hibridación restringentes" se refiere a las condiciones en las que una sonda se hibridará a su subsecuencia objetivo, típicamente en una mezcla compleja de ácido nucleico, pero no a otras secuencias. Las condiciones restringentes dependen de la secuencia y serán diferentes en circunstancias diferentes. Las secuencias más largas se hibridan específicamente a temperaturas más altas. Se encuentra una extensa guía sobre la hibridación de ácidos nucleicos en Tijssen, *Techniques in Biochemistry and Molecular Biology—Hybridization with Nucleic Probes*, "Overview of principles of hybridization and the strategy of nucleic acid assays" (1993). En general, 10 las condiciones restringentes se seleccionan para ser de aproximadamente 5°C a 10°C más bajos que el punto de fusión térmica (T_m) para la secuencia específica a un pH de concentración iónica definido. La T_m es la temperatura (a una concentración iónica, pH y concentración nucleica definidos) en donde el 50 % de las sondas complementarias para el objetivo se hibridan a la secuencia objetivo en equilibrio (debido a que las secuencias objetivo están presentes en exceso, a la T_m , el 50 % de las sondas están ocupadas en equilibrio). Las condiciones restringentes serán aquellas en donde la concentración de sal es menor de aproximadamente 1,0 de ion de sodio, típicamente 15 una concentración de aproximadamente 0,01 a 1,0 M de ion de sodio (u otras sales) a un pH de 7,0 a 8,3 y la temperatura es de cuando menos aproximadamente 30°C para sondas más cortas (por ejemplo, de 10 a 50 nucleótidos) y cuando menos de aproximadamente 60°C para sondas largas (por ejemplo, mayores de 50 nucleótidos). Las condiciones restringentes también se pueden lograr con la adición de agentes desestabilizantes, tales como formamida. Para la hibridación selectiva o específica, una señal positiva es cuando menos dos veces el fondo, opcionalmente una hibridación de 10 veces el fondo. Las condiciones de hibridación restringentes de ejemplo pueden ser como sigue: formamida al 50 %, SSC 5x y SDS al 1 %, incubación a 42°C, o SSC 5x, SDS al 1 %, 20 incubación a 65°C, con lavado en SSC 0,2x y SDS al 0,1 % a 65°C.

25 Los ácidos nucleicos que no se hibridan unos a otros según condiciones restringentes, todavía son sustancialmente idénticos si los polipéptidos que los codifican son sustancialmente idénticos. Esto ocurre, por ejemplo, cuando se crea una copia de un ácido nucleico utilizando la degeneración máxima de codones permitida por el código genético. En estos casos, los ácidos nucleicos típicamente se hibridan según condiciones de hibridación moderadamente restringentes. Las "condiciones de hibridación moderadamente restringentes" de ejemplo incluyen una hibridación en un regulador de formamida al 40 %, NaCl 1M, SDS al 1 % a 37°C y un lavado en SSC 1x a 45°C. Una hibridación 30 positiva es cuando menos el doble del fondo. Los expertos en la técnica reconocerán fácilmente que se pueden utilizar condiciones de hibridación y lavado alternativas para proporcionar condiciones de una restricción similar.

Los términos "idéntico" o porcentaje de "identidad", en el contexto de dos o más ácidos nucleicos o secuencias de polipéptidos, se refieren a una o más secuencias o subsecuencias que son iguales o tienen un porcentaje especificado de residuos de aminoácidos o nucleótidos que son iguales (es decir, identidad del 70%, opcionalmente 35 identidad del 75%, del 80%, del 85%, del 90%, o del 95% sobre una región especificada), cuando se comparan y son alineadas para una máxima correspondencia sobre una ventana de comparación, o una región diseñada, medida utilizando los siguientes algoritmos de comparación de secuencias, o mediante alineación manual e inspección visual. Entonces se dice que estas secuencias son "sustancialmente idénticas". Esta definición también se refiere al complemento de una secuencia de pruebas. Opcionalmente, existe identidad sobre una región que sea 40 de cuando menos aproximadamente 25 a aproximadamente 50 aminoácidos o nucleótidos de longitud, u opcionalmente sobre una región que sea de 75 a 100 aminoácidos o nucleótidos de longitud.

45 Para la comparación de secuencias, típicamente una secuencia actúa como una secuencia de referencia, con la que se comparan las secuencias de prueba. Cuando se utiliza un algoritmo de comparación de secuencias, se introducen las secuencias de prueba y de referencia en una computadora, se designan las coordenadas de las subsecuencias, si es necesario y se designan los parámetros del programa de algoritmo de secuencias. Se pueden utilizar los programas del programa por omisión, o alternativamente se pueden designar parámetros. Entonces el algoritmo de comparación de secuencias calcula el porcentaje de identidad secuencias para las secuencias de prueba en relación con la secuencia de referencia, basándose en los parámetros del programa.

50 Una "ventana de comparación", como se utiliza en la presente memoria, incluye la referencia a un segmento de cualquiera del número de posiciones contiguas seleccionadas a partir del grupo que consiste en de 25 a 500, usualmente de aproximadamente 50 a aproximadamente 200, más usualmente de aproximadamente 100 a aproximadamente 150, en donde se puede comparar una secuencia con una secuencia de referencia del mismo número de posiciones contiguas después de que quedan las dos secuencias óptimamente alineadas. Los procedimientos de alineación de secuencias para comparación son bien conocidos en este campo. Las alineaciones 55 óptimas de secuencias para comparación se pueden conducir, por ejemplo, mediante algoritmo de homología local de Smith y Waterman, *Adv. Appl. Math.* 2:482 (1981), mediante el algoritmo de alineación de homología de Needleman y Wunsch, *J. Mol. Biol.* 48:443 (1970), mediante el procedimiento de búsqueda de similitudes de Pearson y Lipman, *Proc. Nat'l. Acad. Sci. EUA* 85:2444 (1988), mediante implementaciones computarizadas de estos algoritmos (GAP, BESTFIT, FASTA y TFASTA en el Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), o mediante alineación manual e inspección visual (véase, por 60

ejemplo, *Current Protocols in Molecular Biology* (Ausubel y colaboradores, editores, 1995, suplemento)).

Un ejemplo de un algoritmo útil es PILEUP. PILEUP crea una alineación de múltiples secuencias a partir de un grupo de secuencias relacionadas, utilizando alineaciones progresivas por pares para mostrar la relación y el porcentaje de identidad de secuencia. También grafica un árbol o dendrograma que muestra las relaciones de formación de racimos utilizadas para crear la alineación. PILEUP utiliza una simplificación del procedimiento de alineación progresiva de Feng y Doolittle, *J. Mol. Evol.* 35:351-360 (1987). El procedimiento utilizado es similar al procedimiento descrito por Higgins y Sharp, *CABIOS* 5:151-153 (1989). El programa puede alinear hasta 300 secuencias, cada una de una longitud máxima de 5.000 nucleótidos o aminoácidos. El procedimiento de múltiples alineaciones empieza con la alineación por pares de las dos secuencias más similares, produciendo un racimo de dos secuencias alineadas. Este racimo se alinea entonces con la siguiente secuencia o racimo de secuencias alineadas más relacionadas. Dos racimos de secuencias se alinean mediante una extensión simple de la alineación por pares de dos secuencias individuales. La alineación final se logra mediante una serie de alineaciones progresivas por pares. El programa se ejecuta designando las secuencias específicas y sus coordenadas de aminoácidos o nucleótidos para las regiones de comparación de secuencias y designando los parámetros del programa. Utilizando PILEUP, se compara una secuencia de referencia con las otras secuencias de prueba para determinar la relación del porcentaje de identidad de secuencias utilizando los siguientes parámetros: peso de hueco por omisión (3,00), peso de longitud de hueco por omisión (0,10) y huecos de extremo ponderados. PILEUP se puede obtener del paquete de software de análisis de secuencias GCG, por ejemplo versión 7.0 (Devereaux y colaboradores, *Nuc. Acids Res.* 12:387-395 (1984)).

Otro ejemplo de algoritmo que es adecuado para determinar el porcentaje de identidad de secuencias y la similitud de secuencias, es el de los algoritmos BLAST y BLAST 2.0, que se describen en Altschul y colaboradores, *Nuc. Acids Res.* 25:3389-3402 (1977) y en Altschul y colaboradores, *J. Mol. Biol.* 215:403-410 (1990), respectivamente. El software para llevar a cabo los análisis BLAST está públicamente disponible a través del National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>). Este algoritmo involucra identificar primeramente los pares de secuencias de alto puntaje (HSPs), mediante la identificación de palabras cortas de longitud W en la secuencia pedida, que concuerdan o satisfacen algún puntaje umbral de valor positivo T al alinearse con una palabra de la misma longitud en una secuencia de una base de datos. T es referido como el umbral de puntaje de palabra vecina (Altschul y colaboradores, *supra*). Estos impactos de palabra vecina iniciales actúan como semillas para iniciar las búsquedas con el fin de encontrar HSPs más largos que las contengan. Los impactos de palabra se extienden en ambas direcciones a lo largo de cada secuencia por tanto como se pueda aumentar el puntaje de alineación acumulativo. Los puntajes acumulativos se calculan utilizando, para las secuencias de nucleótidos, los parámetros M (puntaje de recompensa para un par de residuos emparejados; siempre >0) y N (puntaje de penalidad para los residuos mal emparejados; siempre <0). Para las secuencias de aminoácidos, se utiliza una matriz de puntaje para calcular el puntaje acumulativo. La extensión de los impactos de palabra en cada dirección se detiene cuando: el puntaje de alineación acumulativo cae fuera por la cantidad X desde su máximo valor alcanzado; el puntaje acumulativo llega hasta cero o menos, debido a la acumulación de una o más alineaciones de residuos de puntaje negativo; o se llega al final de cualquier secuencia. Los parámetros W, T y X del algoritmo BLAST determinan la sensibilidad y velocidad de la alineación. El programa BLASTN (para secuencias de nucleótidos) utiliza por omisión una longitud de palabra (W) de 11, una expectativa (E) de 10, M=5, N=-4 y una comparación de ambas cadenas. Para las secuencias de aminoácidos, el programa BLASTP utiliza por omisión una longitud de palabra de 3 y una expectativa (E) de 10 y la matriz de puntaje BLOSUM62 (véase Henikoff y Henikoff, *Proc. Natl. Acad. Sci. EUA* 89:10915 (1989)), alineaciones (B) de 50, expectativa (E) de 10, M=5, N=-4 y una comparación de ambas cadenas.

El algoritmo BLAST también lleva a cabo un análisis estadístico de la similitud entre dos secuencias (véase, por ejemplo, Karlin y Altschul, *Proc. Natl. Acad. Sci. EUA* 90:5873-5787 (1993)). Una medida de similitud proporcionada por el algoritmo BLAST es la probabilidad de suma más pequeña (P(N)), que proporciona una indicación de la probabilidad por la que se presentaría un emparejamiento entre dos secuencias de nucleótidos o de aminoácidos por azar. Por ejemplo, un ácido nucleico se considera similar a una secuencia de referencia si la probabilidad de suma más pequeña en una comparación de ácido nucleico de prueba con el ácido nucleico de referencia es menor de aproximadamente 0,2, más preferiblemente menor de 0,01 y de una manera muy preferible menor de aproximadamente 0,001.

Composiciones de polinucleotidos

Como se utilizan en la presente memoria, los términos "segmento de ADN" y "polinucleótido" se refieren a una molécula de ADN que se ha aislado para liberarse del ADN genómico total de una especie particular. Por consiguiente, un segmento de ADN que codifica un polipéptido se refiere a un segmento de ADN que contiene una o más secuencias de codificación y no obstante, está sustancialmente aislado lejos de, o purificado libre de, el ADN genómico total de la especie a partir del que se obtiene el segmento de ADN. Dentro de los términos "segmento de ADN" y "polinucleótido", se incluyen los segmentos de ADN y los fragmentos más pequeños de estos segmentos y también los vectores recombinantes, incluyendo, por ejemplo, plásmidos, cósmidos, fagémidos, fagos, virus y similares.

Como será entendido por los expertos en la materia, los segmentos de ADN pueden incluir secuencias genómicas, secuencias extra-genómicas y codificadas por plásmidos y segmentos genéticos diseñados más pequeños que expresen, o se puedan adaptar para expresar, proteínas, polipéptidos, péptidos y similares. Estos segmentos se pueden aislar naturalmente, o pueden ser modificados sintéticamente por la mano del hombre.

- 5 “Aislado”, como se utiliza en la presente memoria, significa que un polinucleótido está sustancialmente lejos de otras secuencias de codificación y que el segmento de ADN no contiene porciones grandes de ADN codificante no relacionado, tales como los fragmentos cromosómicos grandes u otros genes funcionales o regiones codificantes de polipéptidos. Por supuesto, esto se refiere al segmento de ADN como se aisló originalmente y no excluye a los genes o a las regiones codificantes agregadas posteriormente al segmento por la mano del hombre.
- 10 Como será reconocido por el experto, los polinucleótidos pueden ser de una sola cadena (codificante o anti-sentido) o de doble cadena y pueden ser moléculas de ADN (genómico, ADNc, o sintético) o de ARN. Las moléculas de ARN incluyen moléculas de HnARN, que contienen intrones y corresponden a una molécula de ADN de una manera de uno a uno y moléculas de ARNm, que no contienen intrones. Puede haber secuencias codificantes o no codificantes adicionales, pero no es necesario, presentes dentro de un polinucleótido de la presente invención y un polinucleótido puede, pero no necesita, estar enlazado a otras moléculas y/o materiales de soporte.

15 Los polinucleótidos pueden comprender una secuencia nativa (es decir, una secuencia endógena que codifique un antígeno de *Mycobacterium* o una porción del mismo), o pueden comprender una variante, o un equivalente funcional biológico o antigénico de esta secuencia. Las variantes de polinucleótidos pueden contener una o más sustituciones, adiciones, supresiones, y/o inserciones, como se describe adicionalmente más adelante, de preferencia de tal manera que no se disminuya la inmunogenicidad del polipéptido codificado en relación con una proteína tumoral nativa. El efecto sobre la inmunogenicidad del polipéptido codificado se puede evaluar en general como se describe en la presente memoria. El término “variantes” también abarca los genes homólogos de origen xenogénico.

20 Los polinucleótidos de la presente invención, independientemente de la longitud de la secuencia de codificación misma, se pueden combinar con otras secuencias de ADN, tales como promotores, señales de poliadenilación, sitios de enzimas de restricción adicionales, sitios de clonación múltiple, otros segmentos codificantes y similares, de tal manera que su longitud global puede variar considerablemente. Por consiguiente, se contempla que se puede emplear un fragmento de ácido nucleico de casi cualquier longitud, siendo la longitud total de preferencia limitada por la facilidad de preparación y uso en el protocolo de ADN recombinante pretendido.

25 Más aún, será apreciado por los expertos en este campo que, como un resultado de la degeneración del código genético, puede haber muchas secuencias de nucleótidos que codifiquen un polipéptido como se describe en la presente memoria. Algunos de estos polinucleótidos tienen una homología mínima con la secuencia de nucleótidos de cualquier gen nativo. No obstante, la presente invención contempla específicamente los polinucleótidos que varíen debido a las diferencias en el uso de codones, por ejemplo, los polinucleótidos que se optimicen para la selección de codones humanos y/o de primates.

Identificación y caracterización de polinucleótidos

30 Los polinucleótidos se pueden identificar, preparar, y/o manipular empleando cualquiera de una variedad de técnicas bien establecidas. Por ejemplo, un polinucleótido se puede identificar, como se describe con mayor detalle más adelante, mediante el rastreo de un microarreglo de ADNcs para buscar la expresión asociada con tumor (es decir, la expresión que es cuando menos dos veces mayor en un tumor que en el tejido mayor, determinado utilizando un ensayo representativo proporcionado en la presente memoria). Estos rastreos se pueden llevar a cabo, por ejemplo, utilizando un microarreglo de Synteni (Palo Alto, CA) de acuerdo con las instrucciones del fabricante (y esencialmente como es descrito por Schena y colaboradores, *Proc. Natl. Acad. Sci. EUA* 93:10614-10619 (1996), y Heller y colaboradores, *Proc. Natl. Acad. Sci. EUA* 94:2150-2155 (1997)). De una manera alternativa, los polinucleótidos se pueden amplificar a partir del ADNc preparado a partir de células que expresen las proteínas descritas en la presente memoria, tales como las células de *M. tuberculosis*. Estos polinucleótidos se pueden amplificar mediante reacción en cadena de la polimerasa (PCR). Para este planteamiento, se pueden diseñar cebadores específicos de la secuencia, basándose en las secuencias proporcionadas en la presente memoria y se pueden comprar o sintetizar.

40 Se puede utilizar una porción amplificada de un polinucleótido para aislar un gen de longitud completa a partir de una biblioteca adecuada (por ejemplo, una biblioteca de ADNc de *M. tuberculosis*), empleando técnicas bien conocidas. Dentro de estas técnicas, se rastrea una biblioteca (ADNc o genómico) utilizando una o más sondas de polinucleótido o cebadores adecuados para la amplificación. De preferencia, se selecciona por tamaños una biblioteca para incluir moléculas más grandes. También se pueden preferir las bibliotecas cebadas al azar para identificar las regiones 5' y corriente arriba de los genes. Se prefieren las bibliotecas genómicas para obtener intrones y extender las secuencias 5'.

Para las técnicas de hibridación, una secuencia parcial se puede marcar (por ejemplo, mediante traducción de

apriete o marcado de extremo con ³²P) empleando técnicas bien conocidas. Entonces se rastrea en general una biblioteca bacteriana o de bacteriófagos mediante filtros de hibridación que contienen colonias bacterianas desnaturalizadas (o pistas que contienen placas de fagos) con la sonda marcada (véase Sambrook y colaboradores, *Molecular Cloning: A Laboratory Manual* (2000)). Las colonias o placas que se hibridan se seleccionan y se expanden y se aísla el ADN para un análisis adicional. Los clones de ADNc se pueden analizar para determinar la cantidad de secuencia adicional, por ejemplo, mediante reacción en cadena de la polimerasa, utilizando un cebador a partir de la secuencia parcial y un cebador a partir del vector. Se pueden generar mapas de restricción y secuencias parciales para identificar uno o más clones traslapados. Entonces se puede determinar la secuencia completa empleando técnicas convencionales, que pueden involucrar la generación de una serie de clones de supresión. Luego se pueden ensamblar las secuencias traslapadas resultantes en una sola secuencia contigua. Se puede generar una molécula de ADNc de longitud completa mediante el ligamiento de los fragmentos adecuados, empleando técnicas bien conocidas.

De una manera alternativa, existen numerosas técnicas de amplificación para obtener una secuencia de codificación de longitud completa a partir de una secuencia de ADNc parcial. Dentro de estas técnicas, la amplificación se lleva a cabo en general mediante reacción en cadena de la polimerasa. Se puede emplear cualquiera de una variedad de los kits comercialmente disponibles para llevar a cabo el paso de amplificación. Se pueden diseñar cebadores utilizando, por ejemplo, software bien conocido en la técnica. Los cebadores de preferencia son de 22 a 30 nucleótidos de longitud, tienen un contenido de GC de cuando menos el 50 % y se templan a la secuencia objetivo a temperaturas de aproximadamente 68°C a 72°C. La región amplificada se puede secuenciar como se describe anteriormente y se ensamblan las secuencias traslapadas en una secuencia contigua.

Una de estas técnicas de amplificación es la PCR inversa (véase Triglia y colaboradores, *Nucl. Acids Res.* 16:8186 (1988)), que utiliza enzimas de restricción para generar un fragmento en la región conocida del gen. Entonces se circulariza el fragmento mediante ligamiento intramolecular y se utiliza como una plantilla para la reacción en cadena de la polimerasa con cebadores divergentes derivados a partir de la región conocida. Dentro de un planteamiento alternativo, las secuencias adyacentes a una secuencia parcial se pueden recuperar mediante amplificación con un cebador para una secuencia enlazadora y un cebador específico para una región conocida. Las secuencias amplificadas típicamente se someten a una segunda ronda de amplificación con el mismo cebador enlazador y un segundo cebador específico para la región conocida. En el documento WO 96/38591, se describe una variación sobre este procedimiento, que emplea dos cebadores que inician la extensión en direcciones opuestas desde la secuencia conocida. Otra de estas técnicas se conoce como "amplificación rápida de los extremos del ADNc" o RACE. Esta técnica involucra el uso de un cebador interno y un cebador externo, que se hibrida a una región poliA o a una secuencia de vector, para identificar las secuencias que sean 5' y 3' de una secuencia conocida. Las técnicas adicionales incluyen reacción en cadena de la polimerasa de captura (Lagerstrom y colaboradores, *PCR Methods Applic.* 1:111-19 (1991)) y reacción en cadena de la polimerasa caminante (Parker y colaboradores, *Nucl. Acids Res.* 19:3055-60 (1991)). También se pueden emplear otros procedimientos que empleen amplificación con el fin de obtener una secuencia de ADNc de longitud completa.

En ciertos casos, es posible obtener una secuencia de ADNc de longitud completa mediante el análisis de las secuencias proporcionadas en una base de datos de marca de secuencia expresada (EST), tal como la que está disponible en GenBank. Las búsquedas de las EST traslapadas se pueden llevar a cabo en general empleando programas bien conocidos (por ejemplo, búsquedas de NCBI BLAST) y estas EST se pueden utilizar para generar una secuencia de longitud completa contigua. También se pueden obtener secuencias de ADN de longitud completa mediante el análisis de los fragmentos genómicos.

Expresión de polinucleótidos en células huésped

Se pueden utilizar secuencias de polinucleótidos que codifican proteínas de fusión en moléculas de ADN recombinante para dirigir la expresión de un polipéptido en células huésped apropiadas. Debido a la degeneración inherente del código genético, se pueden producir otras secuencias de ADN que codifiquen sustancialmente la misma secuencia de aminoácidos o una secuencia de aminoácidos funcionalmente equivalente y estas secuencias se pueden utilizar para clonar y expresar un polipéptido dado.

Como será entendido por los expertos en la materia, puede ser conveniente, en algunos casos, producir secuencias de nucleótidos que codifiquen polipéptidos que posean codones que no se presenten naturalmente. Por ejemplo, se pueden seleccionar los codones preferidos por un huésped procarionte o eucariótico particular para aumentar el índice de expresión de proteína, o para producir una transcripción de ARN recombinante que tenga las propiedades deseables, tales como una semivida que sea más larga que aquella de una transcripción generada a partir de la secuencia que se presente naturalmente.

Más aún, las secuencias de polinucleótidos de la presente invención se pueden diseñar empleando procedimientos generalmente conocidos en la técnica, con el objeto de alterar las secuencias de codificación de polipéptidos por una variedad de razones, incluyendo, pero no limitándose a, alteraciones que modifiquen la clonación, procesamiento, y/o expresión del producto genético. Por ejemplo, se puede emplear mezcla de ADN mediante fragmentación

aleatoria y reensamble de fragmentos de genes y oligonucleótidos sintéticos mediante reacción en cadena de la polimerasa, para diseñar secuencias de nucleótidos. Además, se puede emplear mutagénesis dirigida al sitio para insertar nuevos sitios de restricción, alterar los patrones de glucosilación, cambiar la preferencia de codones, producir variantes de empalme, o introducir mutaciones, etc.

- 5 Se pueden ligar secuencias de ácidos nucleicos naturales, modificadas, o recombinantes a una secuencia heteróloga, para codificar una proteína de fusión. Por ejemplo, con el fin de rastrear las bibliotecas de péptidos para determinar los inhibidores de la actividad del polipéptido, puede ser útil codificar una proteína quimérica que pueda ser reconocida por un anticuerpo comercialmente disponible. También se puede diseñar una proteína de fusión para que contenga un sitio de disociación localizado entre la secuencia de codificación del polipéptido y la secuencia de proteína heteróloga, de tal manera que el polipéptido se pueda disociar y purificar a partir de la fracción heteróloga.

10 Se pueden sintetizar secuencias que codifiquen un polipéptido deseado, del todo o en parte, empleando procedimientos químicos bien conocidos en la materia (véase Caruthers, M. H. y colaboradores, *Nucl. Acids Res. Symp. Ser.* páginas 215-223 (1980), Horn y colaboradores, *Nucl. Acids Res. Symp. Ser.* páginas 225-232 (1980)). De una manera alternativa, la proteína misma se puede producir empleando procedimientos químicos para sintetizar la secuencia de aminoácidos de un polipéptido, o una porción de la misma. Por ejemplo, se puede llevar a cabo síntesis de péptidos empleando diferentes técnicas en fase sólida (Roberge y colaboradores, *Science* 269:202-204 (1995)) y se puede lograr la síntesis automatizada, por ejemplo, empleando el Sintetizador de Péptidos ABI 431A (Perkin Elmer, Palo Alto, CA).

20 Un péptido recién sintetizado se puede purificar sustancialmente mediante cromatografía de líquidos de alto rendimiento de preparación (por ejemplo, Creighton, *Proteins, Structures and Molecular Principles* (1983)), u otras técnicas comparables disponibles en este campo. La composición de los péptidos sintéticos se puede confirmar mediante análisis de aminoácidos o secuenciación (por ejemplo, el procedimiento de degradación de Edman). Adicionalmente, la secuencia de aminoácidos de un polipéptido, o cualquier parte de la misma, se puede alterar durante la síntesis directa, y/o se puede combinar empleando procedimientos químicos con secuencias de otras proteínas, o cualquier parte de las mismas, para producir un polipéptido variante.

25 Con el objeto de expresar un polipéptido deseado, las secuencias de nucleótidos que codifiquen al polipéptido, o sus equivalentes funcionales, se pueden insertar en el vector de expresión apropiado, es decir, un vector que contenga los elementos necesarios para la transcripción y traducción de la secuencia de codificación insertada. Se pueden emplear los procedimientos que son bien conocidos por los expertos en la técnica para construir vectores de expresión que contengan secuencias que codifiquen un polipéptido de interés y elementos de control de transcripción y de traducción apropiados. Estos procedimientos incluyen las técnicas de ADN recombinante *in vitro*, técnicas sintéticas y recombinación genética *in vivo*. Estas técnicas se describen en Sambrook y colaboradores, *Molecular Cloning, A Laboratory Manual* (2000) y Ausubel y colaboradores, *Current Protocols in Molecular Biology* (actualizado anualmente).

30 Se puede utilizar una variedad de vectores de expresión/ sistemas huésped para contener y expresar las secuencias de polinucleótidos. Estos incluyen, pero no se limitan a, microorganismos, tales como bacterias transformadas con vectores de expresión de ADN recombinante de bacteriófagos, plásmidos, o cósmidos, levadura transformada con vectores de expresión de levadura; sistemas de células de insecto infectados con vectores de expresión de virus (por ejemplo, baculovirus); sistemas de células de plantas transformados con vectores de expresión de virus (por ejemplo, virus de mosaico de coliflor, CaMV; virus de mosaico de tabaco; TMV), o con vectores de expresión bacterianos (por ejemplo, los plásmidos Ti o pBR322); o sistemas de células de animales.

35 Los "elementos de control" o las "secuencias reguladoras" presentes en un sistema de expresión, son las regiones no traducidas del vector - potenciadores, promotores, regiones no traducidas 5' y 3' - que interactúan con las proteínas celulares huésped para llevar a cabo la transcripción y la traducción. Estos elementos pueden variar en su fuerza y especificidad. Dependiendo del sistema de vector y del huésped utilizado, se puede utilizar cualquier número de elementos de transcripción y de traducción adecuados, incluyendo promotores constitutivos e inducibles. Por ejemplo, cuando se clona en sistemas bacterianos, se pueden utilizar promotores inducibles, tales como el promotor lacZ híbrido del fagémido PBLUESCRIPT (Stratagene, La Jolla, Calif.), o el plásmido PSPORT1 (Gibco BRL, Gaithersburg, MD) y similares. En los sistemas celulares de mamífero, en general se prefieren los promotores a partir de genes de mamífero o a partir de virus de mamífero. Si es necesario generar una línea celular que contenga múltiples copias de la secuencia que codifique un polipéptido, convenientemente se pueden utilizar vectores basados en SV40 o EBV, con un marcador seleccionable apropiado.

40 En los sistemas bacterianos, se puede seleccionar un número de vectores de expresión dependiendo del uso pretendido para el polipéptido expresado. Por ejemplo, cuando se necesitan grandes cantidades, por ejemplo para la inducción de anticuerpos, se pueden utilizar vectores que dirijan una expresión de alto nivel de proteínas de fusión que se purifiquen fácilmente. Estos vectores incluyen, pero no se limitan a, los vectores de clonación y expresión en *E. coli* multifuncionales, tales como BLUESCRIPT (Stratagene), en donde la secuencia que codifica el polipéptido de interés se puede ligar en el vector dentro del marco con secuencias para el Met amino-terminal y los siguientes 7

residuos de β -galactosidasa, de tal manera que se produce una proteína híbrida; vectores pIN ((Van Heeke y Schuster, *J. Biol. Chem.* 264:5503-5509 (1989)); y similares. También se pueden utilizar los vectores pGEX (Promega, Madison, Wis.) para expresar polipéptidos extraños como proteínas de fusión con S-transferasa de glutatión (GST). En general, estas proteínas de fusión son solubles y se pueden purificar fácilmente a partir de las células lisadas mediante adsorción en perlas de glutatión-agarosa, seguido por elución en la presencia de glutatión libre. Las proteínas hechas en estos sistemas se pueden diseñar para incluir sitios de disociación de heparina, trombina, o proteasa de factor XA, de tal manera que se pueda liberar el polipéptido clonado de interés a partir de la fracción GST al gusto.

En la levadura, *Saccharomyces cerevisiae*, se pueden utilizar un número de vectores que contengan promotores constitutivos o inducibles, tales como factor alfa, alcohol, oxidasa y PGH. Para las revisiones, véase Ausubel y colaboradores (*supra*) y Grant y colaboradores, *Methods Enzymol.* 153:516-544 (1987).

En los casos en donde se utilizan vectores de expresión en plantas, la expresión de las secuencias que codifiquen los polipéptidos se puede impulsar mediante cualquiera de un número de promotores. Por ejemplo, se pueden utilizar promotores virales, tales como los promotores 35S y 19S de CaMV solos o en combinación con la secuencia líder omega a partir de TMV (Takamatsu, *EMBO J.* 6:307-311 (1987)). De una manera alternativa, se pueden utilizar promotores de plantas, tales como la subunidad pequeña de RUBISCO, o los promotores de choque por calor (Coruzzi y colaboradores, *EMBO J.* 3:1671-1680 (1984); Broglie y colaboradores, *Science* 224:838-843 (1984); y Winter y colaboradores, *Results Probl. Cell Differ.* 17:85-105 (1991)). Estas construcciones se pueden introducir en células de plantas mediante la transformación del ADN directa, o mediante transfección mediada por patógeno. Estas técnicas se describen en un número de revisiones generalmente disponibles (véase, por ejemplo, Hobbs, en *McGraw Hill Yearbook of Science and Technology* páginas 191-196 (1992)).

También se puede utilizar un sistema de insecto para expresar un polipéptido de interés. Por ejemplo, en uno de estos sistemas, se utiliza el virus de polihedrosis nuclear de *Autographa californica* (AcNPV) como un vector para expresar genes extraños en células de *Spodoptera frugiperda* o en *Trichoplusia larvae*. Las secuencias que codifiquen el polipéptido se pueden clonar en la región no esencial del virus, tal como el gen de polihedrina y se someten al control del promotor de polihedrina. La inserción con éxito de la secuencia de codificación del polipéptido hará que el gen de polihedrina se inactive y produzca el virus recombinante careciendo de proteína de recubrimiento. Entonces se pueden utilizar los virus recombinantes para infectar, por ejemplo, células de *S. frugiperda* o *Trichoplusia larvae*, en donde se puede expresar el polipéptido de interés (Engelhard y colaboradores, *Proc. Natl. Acad. Sci. EUA* 91:3224-3227 (1994)).

En las células huésped de mamífero, en general hay un número de sistemas de expresión basados en virus disponibles. Por ejemplo, en los casos en donde se utiliza un adenovirus como un vector de expresión, se pueden ligar las secuencias que codifiquen un polipéptido de interés en un complejo de transcripción/traducción de adenovirus consistente en el promotor tardío y la secuencia líder tripartita. Se puede emplear la inserción en una región E1 o E3 no esencial del genoma viral para obtener un virus viable, que es capaz de expresar el polipéptido en las células huésped infectadas (Logan y Shenk, *Proc. Natl. Acad. Sci. EUA.* 81:3655-3659 (1984)). Además, se pueden utilizar potenciadores de transcripción, tales como el potenciador de virus de sarcoma de Rous (RSV), para aumentar la expresión en células huésped de mamífero. Los procedimientos y protocolos para trabajar con vectores de adenovirus se revisan en Wold, *Adenovirus Methods and Protocols*, 1998. Las referencias adicionales con respecto al uso de vectores de adenovirus se pueden encontrar en *Adenovirus: A Medical Dictionary, Bibliography, and Annotated Research Guide to Internet References*, 2004.

También se pueden utilizar señales de inicio específicas para lograr una traducción más eficiente de las secuencias que codifiquen un polipéptido de interés. Estas señales incluyen el codón de inicio de ATG y las secuencias adyacentes. En los casos en donde se inserten secuencias que codifiquen el polipéptido, su codón de inicio y las secuencias corriente arriba, en el vector de expresión apropiado, pueden no necesitarse señales de control de transcripción o de traducción adicionales. Sin embargo, en los casos en donde solamente se inserte la secuencia de codificación, o una porción de la misma, se deben proporcionar señales de control de traducción exógenas, incluyendo el codón de inicio de ATG. Adicionalmente, el codón de inicio debe estar en el marco de lectura correcto para asegurar la traducción del inserto entero. Los elementos de traducción exógenos y los codones de inicio pueden ser de diferentes orígenes, tanto naturales como sintéticos. La eficiencia de la expresión se puede mejorar mediante la inclusión de potenciadores que sean apropiados para el sistema celular particular que se utilice, tales como los descritos en la bibliografía (Scharf y colaboradores, *Results Probl. Cell Differ.* 20:125-162 (1994)).

Además, se puede seleccionar una cepa de células huésped por su capacidad para modular la expresión de las secuencias insertadas, o para procesar la proteína expresada en la forma deseada. Estas modificaciones del polipéptido incluyen, pero no se limitan a, acetilación, carboxilación, glucosilación, fosforilación, lipidación y acilación. También se puede emplear el procesamiento posterior a la traducción que disocie una forma "prepro" de la proteína, para facilitar la inserción correcta, el pliegue, y/o la función. Se pueden seleccionar diferentes células huésped, tales como CHO, HeLa, MDCK, HEK293 y WI38, que tienen una maquinaria celular específica y los mecanismos característicos para estas actividades posteriores a la traducción, con el fin de asegurar la modificación correcta y el

procesamiento de la proteína extraña.

Para la producción de alto rendimiento a largo plazo de proteínas recombinantes, en general se prefiere la expresión estable. Por ejemplo, las líneas celulares que expresen establemente un polinucleótido de interés, se pueden transformar utilizando vectores de expresión, que pueden contener orígenes de réplica virales y/o elementos de expresión endógenos y un gen marcador seleccionable, sobre el mismo vector o sobre un vector separado. A continuación de la introducción del vector, se puede permitir que las células crezcan durante 1 a 2 días en un medio enriquecido antes de cambiarse al medio selectivo. El propósito del marcador seleccionable es conferir resistencia a la selección y su presencia permite el crecimiento y la recuperación de las células que expresen con éxito las secuencias introducidas. Los clones resistentes de las células establemente transformadas se pueden proliferar empleando técnicas de cultivo de tejido apropiadas para el tipo de célula.

Se puede utilizar cualquier número de sistemas de selección para recuperar las líneas celulares transformadas. Estas incluyen, pero no se limitan a, los genes de cinasa de timidina del virus de herpes simple (Wigler y colaboradores, *Cell* 11:223-32 (1977)) y de fosfo-ribosil-transferasa de adenina (Lowy y colaboradores, *Cell* 22:817-23 (1990)), que se pueden emplear en células tk.sup. o aprt.sup., respectivamente. También, se puede utilizar la resistencia a antimetabolitos, a antibióticos, o a herbicidas como la base para la selección; por ejemplo, dhfr que confiere resistencia al metotrexato (Wigler y colaboradores, *Proc. Natl. Acad. Sci. EUA.* 77:3567-70 (1980)); npt, que confiere resistencia a los aminoglicósidos, neomicina y G-418 (Colbere-Garapin y colaboradores, *J. Mol. Biol.* 150:1-14 (1981)); y als o pat, que confieren resistencia al clorosulfurón y a la acetil-transferasa de fosfinotricina, respectivamente (Murry, *supra*). Se han descrito genes seleccionables adicionales, por ejemplo, trpB, que permite que las células utilicen indol en lugar de triptófano, o hisD, que permite que las células utilicen histinol en lugar de histidina (Hartman y Mulligan, *Proc. Natl. Acad. Sci. EUA.* 85:8047-51 (1988)). Recientemente, el uso de marcadores visibles ha ganado popularidad con marcadores tales como antocianinas, β -glucuronidasa y su sustrato GUS y luciferasa y su sustrato luciferina, que se utilizan ampliamente no solo para identificar los transformantes, sino también para cuantificar la cantidad de expresión de proteína transitoria o estable atribuible a un sistema de vector específico (Rhodes y colaboradores, *Methods Mol. Biol.* 55:121-131 (1995)).

Aunque la presencia/ausencia de la expresión del gen marcador sugiere que también está presente el gen de interés, se puede necesitar confirmar su presencia y su expresión. Por ejemplo, si la secuencia que codifique un polipéptido se inserta dentro de una secuencia del gen marcador, se pueden identificar las células recombinantes que contengan las secuencias por la ausencia de la función del gen marcador. De una manera alternativa, se puede colocar un gen marcador en fila con una secuencia que codifique al polipéptido sometido al control un solo promotor. La expresión del gen marcador en respuesta a la inducción o a la selección, normalmente indica la expresión del gen en fila también. De una manera alternativa, se pueden identificar células huésped que contengan y expresen 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 de ADN-ADN o de ADN-ARN y técnicas de bioensayo o inmunoensayo de proteínas, que incluyen las tecnologías basadas en membrana, en solución, o en chip, para la detección y/o cuantificación del ácido nucleico o de la proteína.

En la técnica se conocen una variedad de protocolos para detectar y medir la expresión de los productos codificados por los polinucleótidos, utilizando anticuerpos policlonales o monoclonales específicos para el producto. Los ejemplos incluyen el ensayo inmunosorbente enlazado con enzima (ELISA), radioinmunoensayo (RIA) y selección de células activada por fluorescencia (FACS). Para algunas aplicaciones, se puede preferir un inmunoensayo basado en monoclonal, de dos sitios, que utilice anticuerpos monoclonales que reaccionen con dos epítomos no interferentes sobre un polipéptido dado, pero también se puede emplear un ensayo de enlace competitivo. Estos y otros ensayos se describen, entre otros lugares, en Hampton y colaboradores, *Serological Methods, a Laboratory Manual* (1990) y en Maddox y colaboradores, *J. Exp. Med.* 158:1211-1216 (1983).

Los expertos en este campo conocen una amplia variedad de marcas y técnicas de conjugación y se pueden utilizar en diferentes ensayos de ácidos nucleicos y aminoácidos. Los medios para producir hibridación marcada o sondas de reacción en cadena de la polimerasa para detectar secuencias relacionadas con los polinucleótidos incluyen oligomarcado, traducción de apriete, marcado de extremo, o amplificación con reacción en cadena de la polimerasa utilizando un nucleótido marcado. De una manera alternativa, las secuencias, o cualesquiera porciones de las mismas, se pueden clonar en un vector para la producción de una sonda de ARNm. Estos vectores se conocen en la técnica, están comercialmente disponibles y se pueden utilizar para sintetizar sondas de ARN *in vitro*, mediante la adición de una polimerasa de ARN apropiada, tal como T7, T3, o SP6 y nucleótidos marcados. Estos procedimientos se pueden conducir empleando una variedad de kits comercialmente disponibles. Las moléculas reporteras o marcas adecuadas que se pueden utilizar incluyen radionúclidos, enzimas, agentes fluorescentes, quimiluminescentes, o cromogénicos, así como sustratos, co-factores, inhibidores, partículas magnéticas y similares.

Las células huésped transformadas con una secuencia de polinucleótido de interés se pueden cultivar según condiciones adecuadas para la expresión y recuperación de la proteína a partir del cultivo celular. La proteína producida por una célula recombinante puede ser secretada o contenida intracelularmente, dependiendo de la secuencia y/o del vector utilizado. Como será entendido por los expertos en la materia, se pueden diseñar vectores

de expresión que contengan los polinucleótidos de la invención para contener secuencias de señales que dirijan la secreción del polipéptido codificado a través de una membrana celular procariótica o eucariótica. Se pueden utilizar otras construcciones recombinantes para unir las secuencias que codifiquen un polipéptido de interés a la secuencia de nucleótidos que codifique un dominio de polipéptido que facilite la purificación de las proteínas solubles. Estos dominios que facilitan la purificación incluyen, pero no se limitan a, péptidos quelantes de metales, tales como los módulos de histidina-triptófano, que permiten la purificación sobre metales inmovilizados, los dominios de la proteína A que permiten la purificación sobre la inmunoglobulina inmovilizada y el dominio utilizado en el sistema de extensión/purificación por afinidad FLAGS (Immunex Corp., Seattle, Wash.). La inclusión de secuencias enlazadoras disociables, tales como aquellas específicas para el factor XA o la enterocinasa (Invitrogen, San Diego, Calif.) entre el dominio de purificación y el polipéptido codificado, se pueden utilizar para facilitar la purificación. Uno de estos vectores de expresión proporciona la expresión de una proteína de fusión que contiene un polipéptido de interés y un ácido nucleico que codifica seis residuos de histidina precedentes a una tiorredoxina o a un sitio de disociación de enterocinasa. Los residuos de histidina facilitan la purificación sobre IMIAC (cromatografía por afinidad de ion de metal inmovilizado), como se describe en Porath y colaboradores, *Prot. Exp. Purif.* 3:263-281 (1992), mientras que el sitio de disociación de enterocinasa proporciona un medio para purificar el polipéptido deseado a partir de la proteína de fusión. Una discusión de los vectores que contienen proteínas de fusión se proporciona en Kroll y colaboradores, *DNA Cell Biol.* 12:441-453 (1993).

Además de los procedimientos de producción recombinante, los polipéptidos de la invención se pueden producir mediante síntesis directa de péptidos, empleando técnicas en fase sólida (Merrifield, *J. Am. Chem. Soc.* 85:2149-2154 (1963)). La síntesis de proteína se puede llevar a cabo empleando técnicas manuales o mediante automatización. La síntesis automatizada se puede lograr, por ejemplo, utilizando el Sintetizador de Péptidos Applied Biosystems 431A (Perkin Elmer). De una manera alternativa, se pueden sintetizar químicamente diferentes fragmentos por separado y se pueden combinar empleando procedimientos químicos para producir la molécula de longitud completa.

25 **Técnicas de suministro de polinucleótidos *in vivo***

En las realizaciones adicionales, se introducen construcciones genéticas que comprendan uno o más de los polinucleótidos de la invención en las células *in vivo*. Esto se puede lograr empleando cualquiera de una variedad de planteamientos bien conocidos, varios de los cuales se ilustran a continuación para propósitos de ilustración.

1. Adenovirus

Uno de los procedimientos preferidos para la administración *in vivo* de una o más secuencias de ácidos nucleicos involucra el uso de un vector de expresión de adenovirus. "Vector de expresión de adenovirus" pretende incluir las construcciones que contienen secuencias de adenovirus suficientes para: (a) soportar el empaque de la construcción y (b) expresar un polinucleótido que se haya clonado en el mismo en una orientación en sentido o anti-sentido. Desde luego, en el contexto de una construcción anti-sentido, la expresión no requiere que se sintetice el producto genético.

El vector de expresión comprende una forma genéticamente diseñada de un adenovirus. El conocimiento de la organización genética del adenovirus, un virus de ADN de doble cadena, lineal, de 36 kb, permite la sustitución de pedazos grandes de ADN adenoviral con secuencias extrañas de hasta 7 kb (Grunhaus y Horwitz, 1992). En contraste con el retrovirus, la infección adenoviral de las células huésped no da como resultado una integración cromosómica, debido a que el ADN adenoviral puede replicarse de una manera episomal sin una genotoxicidad potencial. También, los adenovirus son estructuralmente estables y no se ha detectado ninguna reconfiguración del genoma después de una extensa amplificación. El adenovirus puede infectar virtualmente a todas las células epiteliales, independientemente de su etapa en el ciclo celular. Hasta ahora, la infección adenoviral parece estar ligada solamente a la enfermedad leve, tal como la enfermedad respiratoria aguda en seres humanos.

El adenovirus es particularmente adecuado para utilizarse como un vector de transferencia de genes, debido a su genoma de tamaño mediano, a su facilidad de manipulación, a su alta titulación, a su amplio intervalo de células objetivo y a su alta infectividad. Ambos extremos del genoma viral contienen repeticiones invertidas (ITR) de 100 a 200 pares de bases, que son elementos *cis* necesarios para la réplica y el empaque del ADN. Las regiones temprana (E) y tardía (L) del genoma contienen diferentes unidades de transcripción, que se dividen mediante el establecimiento de la réplica del ADN viral. La región E1 (E1A y E1B) codifica las proteínas responsables de la regulación de la transcripción del genoma viral y unos cuantos genes celulares. La expresión de la región E2 (E2A y E2B) da como resultado la síntesis de las proteínas para la réplica del ADN viral. Estas proteínas están involucradas en la réplica del ADN, en la expresión tardía del gen y en la desactivación de la célula huésped (Renan, 1990). Los productos de los genes tardíos, incluyendo la mayoría de las proteínas de cápside virales, se expresan solamente después de un procesamiento significativo de una sola transcripción primaria emitida por el promotor tardío mayor (MLP). El MLP (localizado a 16.8 m.u.) es particularmente eficiente durante la fase tardía de la infección y todo el ARNm emitido desde este promotor posee una secuencia líder tripartita-5' (TPL) que lo convierte en el ARNm preferido para la traducción.

En un sistema actual, se genera el adenovirus recombinante a partir de recombinación homóloga entre el vector de lanzamiento y el vector de pro-virus. Debido a la posible recombinación entre dos vectores pro-virales, se puede generar el adenovirus de tipo silvestre a partir de este proceso. Por consiguiente, es crítico aislar un solo clon del virus a partir de una placa individual y examinar su estructura genómica.

5 La generación y propagación de los vectores de adenovirus actuales, que son deficientes en réplica, dependen de una línea celular auxiliar única, designada como 293, que se transformó a partir de células de riñón embrionario humano, mediante fragmentos de ADN Ad5 y que expresa constitutivamente las proteínas E1 (Graham y colaboradores, 1977). Debido a que se puede prescindir de la región E3 del genoma de adenovirus (Jones y Shenk, 1978), los vectores de adenovirus actuales, con la ayuda de las células 293, llevan ADN extraño ya sea en la región E1, en la región D3, o en ambas regiones (Graham y Prevec, 1991). En la naturaleza, el adenovirus puede empaquetar aproximadamente el 105 % del genoma de tipo silvestre (Ghosh-Choudhury y colaboradores, 1987), proporcionando la capacidad para aproximadamente 2 kB extra de ADN. Combinado con las aproximadamente 5,5 kB de ADN que son reemplazables en las regiones E1 y E3, la máxima capacidad del vector de adenovirus actual está debajo de 7,5 kB, o aproximadamente el 15 % de la longitud total del vector. Más del 80 % del genoma viral del adenovirus permanece en la estructura base del vector y es la fuente de la citotoxicidad surgida del vector. También, la deficiencia en réplica del virus con E1 suprimida es incompleta. Por ejemplo, se ha observado una filtración de la expresión del gen viral con los vectores actualmente disponibles en altas multiplicidades de infección (MOI) (Mulligan, 1993).

20 Las líneas celulares auxiliares se pueden derivar a partir de células humanas, tales como células de riñón embrionario humano, células de músculo, células hematopoyéticas, u otras células mesenquimales o epiteliales embrionarias humanas. De una manera alternativa, las células auxiliares se pueden derivar a partir de las células de otras especies de mamífero que sean permisivas para el adenovirus humano. Estas células incluyen, por ejemplo, las células Vero u otras células mesenquimales o epiteliales embrionarias de mono. Como se mencionó anteriormente, la línea celular auxiliar actualmente preferida es la 293.

25 Recientemente, Racher y colaboradores (1995) dieron a conocer mejores procedimientos para cultivar las células 293 y propagar el adenovirus. En un formato, se cultivan agregados celulares naturales mediante la inoculación de las células individuales en matraces centrífugos siliconizados de 1 litro (Techne, Cambridge, RU) conteniendo de 100 a 200 mililitros del medio. A continuación de la agitación a 40 revoluciones por minuto, se estima la viabilidad celular con azul de tripano. En otro formato, se emplean microportadores Fibra-Cel (Bibby Sterlin, Stone, RU) (5 gramos/litro) como sigue. Se agrega un inóculo celular, resuspendido en 5 mililitros del medio, al portador (50 mililitros) en un matraz Erlenmeyer de 250 mililitros y se deja estacionario, con agitación ocasional, durante 1 a 4 horas. Luego se reemplaza el medio con 50 mililitros de medio fresco y se inicia la agitación. Para la producción del virus, se permite que las células crezcan hasta aproximadamente una confluencia del 80 %, después de cuyo tiempo, se reemplaza el medio (hasta el 25 % del volumen final) y se agrega el adenovirus a una multiplicidad de infección de 0,05. Los cultivos se dejan estacionarios durante la noche, a continuación de lo que, se aumenta el volumen hasta el 100 % y se comienza la agitación durante otras 72 horas.

40 De una forma diferente que no sea el requerimiento de que el vector de adenovirus sea defectuoso en réplica, o cuando menos condicionalmente defectuoso, no se cree que la naturaleza del vector de adenovirus sea crucial para la práctica con éxito de la invención. El adenovirus puede ser de cualquiera de los 42 serotipos conocidos diferentes o subgrupos A-F. El adenovirus tipo 5 o subgrupo C es el material de partida preferido con el objeto de obtener un vector de adenovirus defectuoso en réplica condicional para utilizarse en la presente invención, debido a que el adenovirus tipo 5 es un adenovirus humano acerca del que se conoce una gran cantidad de información bioquímica y genética, e históricamente se ha utilizado para la mayoría de las construcciones que emplean adenovirus como un vector.

45 Como se mencionó anteriormente, el vector típico de acuerdo con la presente invención es defectuoso en réplica y no tendrá una región de adenovirus E1. Por lo tanto, será más conveniente introducir el polinucleótido que codifique el gen de interés en la posición a partir del que se hayan removido las secuencias que codifiquen E1. Sin embargo, la posición de inserción de la construcción dentro de las secuencias de adenovirus no es crítica para la invención. El polinucleótido que codifique el gen de interés también se puede insertar en lugar de la región E3 suprimida, en los vectores de reemplazo de E3, como es descrito por Karlsson y colaboradores (1986), o en la región E4, en donde una línea celular auxiliar o un virus auxiliar complementan al defecto de E4.

55 El adenovirus es fácil de cultivar y manipular y exhibe un amplio rango de huéspedes *in vitro* e *in vivo*. Este grupo de virus se puede obtener en altas titulaciones, por ejemplo, de 10^9 a 10^{11} unidades formadoras de placas por mililitro y son altamente infecciosos. El ciclo de vida del adenovirus no requiere de la integración en el genoma de la célula huésped. Los genes extraños suministrados por los vectores de adenovirus son episomales y por consiguiente, tienen una baja genotoxicidad para las células huésped. No se han reportado efectos secundarios en los estudios de vacunación con adenovirus de tipo silvestre (Couch y colaboradores, 1963; Top y colaboradores, 1971), demostrando su seguridad y su potencial terapéutico como vectores de transferencia genética *in vivo*.

Los vectores de adenovirus se han utilizado en la expresión de genes eucarióticos (Levrero y colaboradores, 1991; Gomez-Foix y colaboradores, 1992) y el desarrollo de vacunas (Grunhaus y Horwitz, 1992; Graham y Prevec, 1992). Recientemente, estudios con animales sugirieron que se podría utilizar el adenovirus recombinante para la terapia genética (Stratford-Perricaudet y Perricaudet, 1991; Stratford-Perricaudet y colaboradores, 1990; Rich y colaboradores, 1993). Los estudios en la administración de adenovirus recombinante a diferentes tejidos incluyen instilación en la tráquea (Rosenfeld y colaboradores, 1991; Rosenfeld y colaboradores, 1992), inyección muscular (Ragot y colaboradores, 1993), inyecciones intravenosas periféricas (Herz y Gerard, 1993), e inoculación estereotáctica en el cerebro (Le Gal La Salle y colaboradores, 1993).

Los vectores de adenovirus se pueden originar a partir de adenovirus humanos. De una manera alternativa, se pueden originar a partir de adenovirus de otras especies, por ejemplo de chimpancé, que pueden tener la ventaja de que los vectores virales no se neutralizan por los anticuerpos contra los adenovirus humanos que circulan en muchos sujetos humanos (véase, por ejemplo, Tatsis N. y colaboradores (2005) *Gene Ther.* diciembre 1; [publicación electrónica antes de impresión]).

2. Retrovirus

Los retrovirus son un grupo de virus de ARN de una sola cadena caracterizados por una capacidad para convertir su ARN en ADN de doble cadena en las células infectadas mediante un proceso de transcripción inversa (Coffin, 1990). Entonces el ADN resultante se integra establemente en los cromosomas celulares como un pro-virus y dirige la síntesis de las proteínas virales. La integración da como resultado la integración de las secuencias genéticas virales en la célula receptora y sus descendientes. El genoma retroviral contiene tres genes, gag, pol y env, que codifican para las proteínas de cápside, la enzima polimerasa y los componentes de envoltura, respectivamente. Una secuencia que se encuentra corriente arriba desde el gen gag, contiene una señal para empaquetar el genoma en viriones. Hay dos secuencias de repetición terminal larga (LTR) presentes en los extremos 5' y 3' del genoma viral. Estas contienen fuertes secuencias promotoras y potenciadoras y también se requieren para la integración en el genoma de la célula huésped (Coffin, 1990).

Con el objeto de construir un vector retroviral, se inserta un ácido nucleico que codifique una o más secuencias de oligonucleótidos o de polinucleótidos de interés, en el genoma viral, en el lugar de ciertas secuencias virales, para producir un virus que sea defectuoso en réplica. Con el objeto de producir viriones, se construye una línea celular de empaque que contenga a los genes gag, pol y env, pero sin la región terminal larga y los componentes de empaque (Mann y colaboradores, 1983). Cuando se introduce un plásmido recombinante que contenga un ADNc, junto con la región germinal larga retroviral y las secuencias de empaque, en esta línea celular (mediante precipitación con fosfato de calcio, por ejemplo), la secuencia de empaque permite que la transcripción del ARN del plásmido recombinante se empaque en las partículas virales, que luego son secretadas hacia el medio de cultivo (Nicolas y Rubenstein, 1988; Temin, 1986; Mann y colaboradores, 1983). Entonces se recolecta el medio que contiene los retrovirus recombinantes, opcionalmente se concentra y se utiliza para la transferencia genética. Los vectores retrovirales son capaces de infectar a una amplia variedad de tipos de células. Sin embargo, la integración y la expresión estable requieren de la división de las células huésped (Paskind y colaboradores, 1975).

Recientemente se desarrolló un planteamiento novedoso diseñado para permitir la dirección específica de los vectores de retrovirus, basado en la modificación química de un retrovirus mediante la adición química de residuos de lactosa a la envoltura viral. Esta modificación podría permitir la infección específica de los hepatocitos por medio de los receptores de sialoglucoproteína.

Se diseñó un planteamiento diferente para dirigir los retrovirus recombinantes, en donde se utilizaron anticuerpos biotinilados contra una proteína de envoltura retroviral y contra un receptor celular específico. Los anticuerpos se acoplaron por medio de los componentes de biotina, utilizando estreptavidina (Roux y colaboradores, 1989). Utilizando los anticuerpos contra los antígenos del complejo de histocompatibilidad mayor clase I y clase II, demostraron la infección de una variedad de células humanas que perforan estos antígenos superficiales con un virus ecotrópico *in vitro* (Roux y colaboradores, 1989).

3. Virus adeno-asociado

AAV (Ridgeway, 1988; Hermonat y Muzycska, 1984) es un parvovirus, descubierto como una contaminación de las administraciones adenovirales. Es un virus ubicuo (los anticuerpos están presentes en el 85 % de la población humana de los EE.UU.) que no se ha vinculado con ninguna enfermedad. También se clasifica como un dependovirus, debido a que sus réplicas dependen de la presencia de un virus auxiliar, tal como adenovirus. Se han aislado cinco serotipos, de los que el AAV-2 es el mejor caracterizado. AAV tiene un ADN lineal de una sola cadena que se encapsida en las proteínas de cápside VP1, VP2 y VP3, para formar un virión icosaédrico de 20 a 24 nanómetros de diámetro (Muzyczka y McLaughlin, 1988).

El ADN de AAV es de aproximadamente 4,7 kilobases de largo. Contiene dos marcos de lectura abierta y está flanqueado por dos ITR. Existen dos genes principales en el genoma AAV: *rep* y *cap*. El gen *rep* codifica para las proteínas responsables de las réplicas virales, mientras que *cap* codifica para la proteína de cápside VP1-3. Cada

ITR forma una estructura de horquilla en forma de T. Estas repeticiones terminales son los únicos componentes *cis* esenciales del AAV para la integración cromosómica. Por consiguiente, el AAV se puede utilizar como un vector con todas las secuencias de codificación viral removidas y reemplazadas por el casete de genes para la administración. Se han identificado tres promotores virales y se denominan como p5, p19 y p40, de acuerdo con su posición en el mapa. La transcripción desde p5 y p19 da como resultado la producción de las proteínas rep y la transcripción desde p40 produce las proteínas de cápside (Hermonat y Muzyczka, 1984).

Existen varios factores que hicieron que los investigadores estudiaran la posibilidad de utilizar rAAV como un vector de expresión. Uno es que los requerimientos para administrar un gen con el fin de integrarlo en el cromosoma huésped son sorprendentemente pocos. Es necesario tener las ITR de 145 pares de bases, que son solamente el 6 % del genoma de AAV. Esto deja espacio en el vector para ensamblar una inserción de ADN de 4,5 kb.

AAV también es una buena elección de vehículos de suministro debido a su seguridad. Hay un mecanismo de rescate relativamente complicado: no solamente el adenovirus de tipo silvestre sino también los genes AAV son requeridos para movilizar rAAV. De la misma manera, AAV no es patógeno y no está asociado con ninguna enfermedad. La remoción de las secuencias de codificación viral minimiza las reacciones inmunes a la expresión del gen viral y por consiguiente, rAAV no provoca una respuesta inflamatoria.

4. Otros vectores virales como construcciones de expresión.

Se pueden emplear otros vectores de virales como construcciones de expresión en la presente invención para la administración de secuencias de oligonucleótidos o polinucleótidos a una célula huésped. Se pueden emplear vectores derivados a partir de virus tales como virus de vacuna (Ridgeway, 1988; Coupar y colaboradores, 1988), lentivirus, virus de polio y virus de herpes. Estos ofrecen varias características atractivas para diferentes células de mamífero (Friedmann, 1989; Ridgeway, 1988; Coupar y colaboradores, 1988; Horwich y colaboradores, 1990).

Con el reciente reconocimiento de los virus de hepatitis B defectuosos, se ganó una nueva perspectiva sobre la relación de estructura-función de diferentes secuencias virales. Los estudios *in vitro* mostraron que el virus podría retener la capacidad para el empaque dependiente del auxiliar y la transcripción inversa, a pesar de la supresión de hasta el 80 % de su genoma (Horwich y colaboradores, 1990). Esto sugirió que grandes porciones del genoma podrían ser reemplazadas con el material genético extraño. El hepatotropismo y la persistencia (integración) fueron propiedades particularmente atractivas para la transferencia del gen dirigida al hígado. Chang y colaboradores (1991) introdujeron el gen de acetil-transferasa de cloranfenicol (CAT) en el genoma del virus de hepatitis B de pato en lugar de las secuencias de codificación de polimerasa, superficiales y pre-superficiales. Se co-transfectó con el virus de tipo silvestre en una línea celular de hepatoma de aves. Se utilizaron medios de cultivo conteniendo altas titulaciones del virus recombinante para infectar los hepatocitos de pato primarios. La expresión del gen CAT estable se detectó durante cuando menos 24 días después de la transfección (Chang y colaboradores, 1991).

5. Vectores no virales.

Con el objeto de efectuar la expresión de las secuencias de polinucleótidos de la presente invención, se debe administrar la construcción de expresión a una célula. Esta administración se puede llevar a cabo *in vitro*, como en los procedimientos de laboratorio para transformar líneas celulares, o *in vivo* o *ex vivo*, como en el tratamiento de ciertos estados de enfermedad. Como se describe anteriormente, un mecanismo preferido para la administración es mediante infección viral, en donde la construcción de expresión se encapsula en una partícula viral infecciosa.

Una vez que se ha suministrado la construcción de expresión a la célula, se puede colocar el ácido nucleico que codifica las secuencias de oligonucleótidos o polinucleótidos deseadas y se expresa en diferentes sitios. En ciertas realizaciones, el ácido nucleico que codifica la construcción se puede integrar establemente en el genoma de la célula. Esta integración puede estar en la localización y orientación específicas mediante recombinación homóloga (reemplazo de genes), o se puede integrar en una localización no específica aleatoria (aumento de genes). En todavía otras realizaciones, el ácido nucleico se puede mantener establemente en la célula como un segmento episomal separado de ADN. Estos segmentos de ácidos nucleicos o "episomas" codifican secuencias suficientes para permitir el mantenimiento y la réplica, independientemente de, o en sincronización con, el ciclo de la célula huésped. La manera en que se suministra la construcción de expresión a una célula y en dónde permanece el ácido nucleico en la célula, dependen del tipo de construcción de expresión empleada.

En ciertas realizaciones de la invención, la construcción de expresión que comprenda una o más secuencias de oligonucleótidos o polinucleótidos puede simplemente consistir en ADN recombinante desnudo o en plásmidos. La transferencia de la construcción se puede llevar a cabo mediante cualquiera de los procedimientos mencionados anteriormente, que permeabilizan físicamente o químicamente la membrana celular. Esto es particularmente aplicable para la transferencia *in vitro*, pero se puede aplicar también al uso *in vivo*. Dubensky y colaboradores (1984) inyectaron con éxito ADN de poliomavirus en la forma de precipitados en fosfato de calcio en hígado y bazo de ratones adultos y recién nacidos, demostrando la réplica viral activa y una infección aguda. Benvenisty y Reshef (1986) también demostraron que la inyección intraperitoneal directa de plásmidos precipitados con fosfato de calcio, da como resultado la expresión de los genes transfectados. Se prevé que el ADN que codifique un gen de interés

también se puede transferir de una manera similar *in vivo* y expresar el producto genético.

Otra realización de la invención para transferir una construcción de expresión de ADN desnudo a las células puede involucrar el bombardeo de partículas. Este procedimiento depende de la capacidad para acelerar los microproyectiles recubiertos con ADN hasta una alta velocidad, permitiéndoles perforar las membranas celulares y entrar a las células sin matarlas (Klein y colaboradores, 1987). Se han desarrollado varios dispositivos para acelerar partículas pequeñas. Uno de estos dispositivos se apoya en una descarga de alto voltaje para generar una corriente eléctrica, que a su vez proporciona la fuerza motriz (Yang y colaboradores, 1990). Los microproyectiles utilizados han consistido en sustancias biológicamente inertes, tales como perlas de tungsteno o de oro.

Los órganos seleccionados, incluyendo hígado, piel y tejido muscular de ratas y ratones, han sido bombardeados *in vivo* (Yang y colaboradores, 1990; Zelenin y colaboradores, 1991). Esto puede requerir de la exposición quirúrgica del tejido o de las células, para eliminar cualquier tejido que intervenga entre la pistola y el órgano objetivo, es decir, el tratamiento *ex vivo*.

Composiciones de polipéptidos

La presente invención, en otros aspectos, proporciona composiciones de polipéptidos.

En general se pueden identificar porciones inmunógenas empleando técnicas bien conocidas, tales como las que se resumen por Paul, *Fundamental Immunology*, 3ª Edición, 243-247 (1993) y las referencias citadas en el mismo. Estas técnicas incluyen el rastreo de polipéptidos para determinar su capacidad para reaccionar con anticuerpos específicos del antígeno, anti-sueros, y/o líneas de células-T o clones. Como se utilizan en la presente memoria, los anti-sueros y los anticuerpos "específicos del antígeno" si se enlazan específicamente a un antígeno (es decir, reaccionan con la proteína en un ELISA o en otro inmunoensayo y no reaccionan de una manera detectable con las proteínas no relacionadas). Estos anti-sueros y anticuerpos se pueden preparar como se describe en la presente memoria y empleando técnicas bien conocidas. Una porción inmunógena de una proteína de *Mycobacterium sp.*, es una porción que reacciona con estos anti-sueros y/o células-T a un nivel que no es sustancialmente menor que la reactividad del polipéptido de longitud completa (por ejemplo, en un ELISA y/o en un ensayo de reactividad de células-T). Estas porciones inmunógenas pueden reaccionar dentro de estos ensayos a un nivel que es similar a, o mayor que, la reactividad del polipéptido de longitud completa. Estos rastreos se pueden llevar a cabo en general empleando procedimientos bien conocidos por los expertos en este campo, tales como aquellos descritos en Harlow y Lane, *Antibodies: A Laboratory Manual* (1988) y *Using Antibodies: A Laboratory Manual* (1998). Por ejemplo, un polipéptido se puede inmovilizar sobre un soporte sólido y se puede poner en contacto con el suero del paciente para permitir el enlace de los anticuerpos dentro del suero con el polipéptido inmovilizado. Entonces se puede remover el suero no enlazado y se detectan los anticuerpos enlazados utilizados, por ejemplo, Proteína A marcada con ¹²⁵I.

Los polipéptidos se pueden preparar empleando cualquiera de una variedad de técnicas bien conocidas. Los polipéptidos recombinantes codificados por secuencias de ADN, como se describen anteriormente, se pueden preparar fácilmente a partir de las secuencias de ADN empleando cualquiera de una variedad de vectores de expresión conocidos por los expertos en la materia. La expresión se puede lograr en cualquier célula huésped apropiada que se haya transformado o transfectado con un vector de expresión que contenga una molécula de ADN que codifique un polipéptido recombinante. Las células huésped adecuadas incluyen células de procariotas, de levadura y de eucariotas superiores, tales como células de mamífero y células de plantas. De preferencia, las células huésped empleadas son las líneas celulares de *E. coli*, de levadura, o de mamífero, tales como COS o CHO. Los sobrenadantes de los sistemas de huésped/vector adecuados que secreten la proteína o el polipéptido recombinante en el medio de cultivo, se pueden concentrar primeramente utilizando un filtro comercialmente disponible. A continuación de la concentración, el concentrado se puede aplicar a una matriz de purificación adecuada, tal como una matriz de afinidad o una resina de intercambio de iones. Finalmente, se pueden emplear uno o más pasos de HPLC en fase inversa para purificar adicionalmente un polipéptido recombinante.

Los polipéptidos que tengan menos de aproximadamente 100 aminoácidos y en general menos de aproximadamente 50 aminoácidos, también se pueden generar por medios sintéticos, empleando técnicas bien conocidas por los expertos en este campo. Por ejemplo, estos polipéptidos se pueden sintetizar empleando cualquiera de las técnicas en fase sólida comercialmente disponibles, tales como el procedimiento de síntesis en fase sólida de Merrifield, en donde se agregan en secuencia los aminoácidos a una cadena de aminoácidos creciente. Véase Merrifield, *J. Am. Chem. Soc.* 85:2149-2146 (1963). El equipo para la síntesis automatizada de polipéptidos está comercialmente disponible con proveedores tales como Perkin Elmer/Applied BioSystems Division (Foster City, CA) y se puede operar de acuerdo con las instrucciones del fabricante.

Dentro de ciertas realizaciones específicas, un polipéptido puede ser una proteína de fusión que comprenda múltiples polipéptidos, como se describe en la presente memoria, o que comprenda cuando menos un polipéptido como se describe en la presente memoria y una secuencia no relacionada. Un componente de fusión, por ejemplo, puede asistir para proporcionar epítopos auxiliares-T (un componente de fusión inmunológico), de preferencia los

epítomos auxiliares-T reconocidos por los seres humanos, o puede asistir en la expresión de la proteína (un potenciador de expresión) en rendimientos más altos que la proteína recombinante nativa. Ciertos componentes de fusión preferidos son componentes de fusión tanto inmunológicos como potenciadores de la expresión. Se pueden seleccionar otros componentes de fusión para aumentar la solubilidad de la proteína, o para hacer posible que se dirija la proteína hacia los compartimientos intracelulares deseados. Todavía además, los componentes de fusión incluyen marcos de afinidad, que facilitan la purificación de la proteína.

Las proteínas de fusión se pueden preparar en general empleando técnicas convencionales, incluyendo conjugación química. De preferencia, una proteína de fusión se expresa como una proteína recombinante, permitiendo la producción de mayores niveles, en relación con una proteína no fusionada, en un sistema de expresión. Dicho de una manera breve, las secuencias de ADN que codifiquen a los componentes del polipéptido se pueden ensamblar por separado y se pueden ligar en un vector de expresión apropiado. El extremo 3' de la secuencia de ADN que codifique un componente de polipéptido se liga, con o sin un enlazador peptídico, al extremo 5' de una secuencia de ADN que codifique el segundo componente de polipéptido, de tal manera que los marcos de lectura de las secuencias estén en fase. Esto permite la traducción hasta una sola proteína de fusión que retiene la actividad biológica de ambos polipéptidos componentes.

Se puede emplear una secuencia enlazadora peptídica para separar los primero y segundo componentes de polipéptido por una distancia suficiente para asegurar que cada polipéptido se pliegue en sus estructuras secundarias y terciarias. Esta secuencia enlazadora peptídica se incorpora en la proteína de fusión empleando técnicas convencionales bien conocidas en este campo. Las secuencias enlazadoras peptídicas adecuadas se pueden seleccionar basándose en los siguientes factores: (1) su capacidad para adoptar una conformación extendida flexible; (2) su incapacidad para adoptar una estructura secundaria que podría interactuar con los epítomos funcionales sobre los primero y segundo polipéptidos; y (3) la falta de residuos hidrofóbicos o cargados, que podrían reaccionar con los epítomos funcionales del polipéptido. Las secuencias enlazadoras peptídicas preferidas contienen los residuos Gly, Asn y Ser. También se pueden utilizar otros aminoácidos casi neutros, tales como Thr y Ala, en la secuencia enlazadora. Las secuencias de aminoácidos que se pueden emplear útilmente como enlazadoras incluyen aquellas que se dan a conocer por Maratea y colaboradores, *Gene* 40:39-46 (1985); Murphy y colaboradores, *Proc. Natl. Acad. Sci. EUA* 83:8258-8262 (1986); Patente de los EE.UU. n.º: 4.935.233 y Patente de los EE.UU. n.º: 4.751.180. La secuencia enlazadora puede ser en general de 1 a aproximadamente 50 aminoácidos de longitud. Las secuencias enlazadoras no se requieren cuando los primero y segundo polipéptidos tienen regiones de aminoácidos N-terminales no esenciales que se pueden utilizar para separar los dominios funcionales y para prevenir la interferencia estérica.

Las secuencias de ADN ligadas se enlazan operativamente a los elementos reguladores de transcripción o de traducción adecuados. Los elementos reguladores responsables de la expresión del ADN se localizan solamente a 5' para la secuencia de ADN que codifica los primeros polipéptidos. De una manera similar, los codones de paro requeridos para finalizar las señales de terminación de traducción y de transcripción están solamente presentes a 3' para las secuencias de ADN que codifican el segundo polipéptido.

En las realizaciones preferidas, un componente de fusión inmunológico se deriva a partir de la proteína D, una proteína superficial de la bacteria gram-negativa *Haemophilus influenza* B (documento WO 91/18926). De preferencia, un derivado de proteína D comprende aproximadamente la primera tercera parte de la proteína (por ejemplo, los primeros 100 a 110 aminoácidos N-terminales) y un derivado de proteína D puede estar liquidado. Dentro de ciertas realizaciones preferidas, se incluyen los primeros 109 residuos de un componente de fusión de lipoproteína D en el extremo N-terminal para proporcionar al polipéptido epítomos de células-T exógenos adicionales y para aumentar el nivel de expresión en *E. coli* (funcionando de esta manera como un potenciador de expresión). La cola del lípido asegura la presentación óptima del antígeno ante las células presentadoras de antígeno. Otros componentes de fusión incluyen la proteína no estructural del virus de influenza, NS 1 (hemaglutinina). Típicamente, se utilizan los 81 aminoácidos N-terminales, aunque se pueden utilizar diferentes fragmentos que incluyan a los epítomos auxiliares-T.

En otra realización, el componente de fusión inmunológico es una proteína conocida como LYTA, o una porción de la misma (de preferencia una porción C-terminal). LYTA se deriva a partir de *Streptococcus pneumoniae*, que sintetiza una amidasa de N-acetil-L-alanina conocida como amidasa LYTA (codificada por el gen *LytA*; *Gene* 43:265-292 (1986)). LYTA es una autolisina que degrada específicamente ciertos enlaces en la estructura base del peptidoglucano. El dominio C-terminal de la proteína LYTA es responsable de la afinidad con la colina, o con algunos análogos de colina, tales como DEAE. Esta propiedad se ha explotado para el desarrollo de plásmidos que expresan C-LYTA de *E. coli* útiles para la expresión de proteínas de fusión. La purificación de las proteínas híbridas que contienen el fragmento C-LYTA en el término amino ya ha sido descrita (véase *Biotechnology* 10:795-798 (1992)). Dentro de una realización preferida, se puede incorporar una porción de repetición de LYTA en una proteína de fusión. Una porción de repetición se encuentra en la región C-terminal empezando en el residuo 178. Una porción de repetición particularmente preferida incorpora los residuos 188-305.

En general, los polipéptidos (incluyendo las proteínas de fusión) y los polinucleótidos descritos en la presente

memoria, se aíslan. Un polipéptido o polinucleótido “aislado”, es uno que se remueve de su medio ambiente original. Por ejemplo, una proteína que se presenta naturalmente se aísla si se separa de algunos o todos los materiales coexistentes en el sistema natural. De preferencia, estos polipéptidos son cuando menos aproximadamente el 90 % puros, más preferiblemente cuando menos aproximadamente el 95 % puros y de una manera muy preferible cuando menos aproximadamente el 99 % puros. Se considera que un polinucleótido está aislado si, por ejemplo, se clona en un vector que no sea una parte del medio ambiente natural.

Composiciones farmacéuticas

En las realizaciones adicionales, la presente invención se refiere a una formulación de uno o más de los polinucleótidos o polipéptidos dados a conocer en la presente memoria, en soluciones farmacéuticamente aceptables para administrarse a una célula o a un animal, ya sea solos, o bien en combinación con una o más realizaciones diferentes de terapia.

Se entenderá que, si se desea, el segmento de ácido nucleico (por ejemplo, ARN o ADN) que expresa un polipéptido como se da a conocer en la presente memoria, se puede administrar en combinación con otros agentes también, tales como, por ejemplo, otras proteínas o polipéptidos, o diferentes agentes farmacéuticamente activos, incluyendo agentes quimioterapéuticos efectivos contra una infección por *M. tuberculosis*. De hecho, virtualmente no hay límite para otros componentes que también se puedan incluir, dado que los agentes adicionales no provocan un efecto adverso significativo después de su contacto con las células objetivo o los tejidos del huésped. Por consiguiente, las composiciones se pueden administrar junto con otros diferentes agentes, como se requiera en la instancia particular. Estas composiciones se pueden purificar a partir de células huésped o de otras fuentes biológicas, o de una manera alternativa, se pueden sintetizar químicamente como se describe en la presente memoria. De la misma manera, estas composiciones pueden comprender además composiciones de ARN o ADN sustituidas o derivadas.

La formulación de los excipientes y soluciones de vehículo farmacéuticamente aceptables es bien conocida por los expertos en la materia, así como el desarrollo de regímenes de dosificación y tratamiento adecuados para utilizar las composiciones particulares descritas en la presente memoria en una variedad de regímenes de tratamiento, incluyendo, por ejemplo, administración y formulación oral, parenteral, intravenosa, intranasal, e intramuscular. Típicamente, las formulaciones que comprenden una cantidad terapéuticamente efectiva suministran de aproximadamente 2 microgramos a aproximadamente 50 microgramos del polipéptido Mtb72f por administración, más típicamente de aproximadamente 5 microgramos a aproximadamente 40 microgramos del polipéptido Mtb72f por administración.

1. Suministro oral.

En ciertas aplicaciones, las composiciones farmacéuticas dadas a conocer en la presente memoria se pueden administrar mediante administración oral a un animal. Como tales, estas composiciones se pueden formular con un diluyente inerte o con un vehículo comestible asimilable, o se pueden encerrar en cápsulas de gelatina de cubierta dura o blanda, o se pueden comprimir en comprimidos, o se pueden incorporar directamente con el alimento de la dieta.

Los compuestos activos inclusive se pueden incorporar con excipientes y se pueden utilizar en la forma de comprimidos ingeribles, comprimidos bucales, trociscos, cápsulas, elixires, suspensiones, jarabes, obleas y similares (Mathiowitz y colaboradores, 1997; Hwang y colaboradores, 1998; Patente de los EE.UU. n.º: 5.641.515; Patente de los EE.UU. n.º: 5.580.579 y Patente de los EE.UU. n.º: 5.792.451, cada uno específicamente incorporado a la presente memoria por referencia en su totalidad). Los comprimidos, trociscos, píldoras, cápsulas y similares, también pueden contener lo siguiente: un aglutinante, como goma de tragacanto, acacia, almidón de maíz, o gelatina; excipientes, tales como difosfato de calcio; un agente desintegrante, 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 saborizante, tal como menta, aceite de hierbabuena, o saborizante de cereza. Cuando la forma unitaria de dosificación es una cápsula, puede contener, además de los materiales del tipo anterior, un vehículo líquido. Puede haber otros diferentes materiales presentes como recubrimientos o para modificar de otra manera la forma física de la unidad de dosificación. Por ejemplo, los comprimidos, píldoras, o cápsulas se pueden recubrir con goma laca, azúcar, o ambos. Un jarabe de elixir puede contener al compuesto activo con sacarosa como un agente edulcorante, metil- y propil-parabenos como conservadores, un tinte y un saborizante, tal como saborizante de cereza o de naranja. Por supuesto, cualquier material utilizado en la preparación de cualquier forma unitaria de dosificación debe ser farmacéuticamente puro y sustancialmente no tóxico en las cantidades empleadas. Además, los compuestos activos se pueden incorporar en preparaciones y formulaciones de liberación sostenida.

Típicamente, estas formulaciones normalmente contienen entre 2 microgramos y 50 microgramos del polipéptido Mtb72f. Naturalmente, la cantidad de los compuestos activos en cada composición terapéuticamente útil se puede preparar de tal manera que se obtenga una dosificación adecuada en cualquier dosis unitaria dada del compuesto. Los expertos en la técnica de la preparación de estas formulaciones farmacéuticas contemplarán factores tales como

la solubilidad, la biodisponibilidad, la vida de anaquel biológica, la vía de administración, la vida de anaquel del producto, así como otras consideraciones farmacológicas y como tales, pueden ser deseables una variedad de dosificación y regímenes de tratamiento.

5 Para la administración oral, las composiciones de la presente invención se pueden incorporar de una manera alternativa con uno o más excipientes en la forma de un enjuague bucal, dentífrico, comprimido bucal, aerosol oral, o formulación sublingual oralmente administrada. Por ejemplo, un enjuague bucal se puede preparar incorporando el ingrediente activo en la cantidad requerida en un solvente apropiado, tal como una solución de borato de sodio (Solución de Dobell). De una manera alternativa, el ingrediente activo se puede incorporar en una solución oral, tal como una que contenga borato de sodio, glicerina y bicarbonato de potasio, o se puede dispersar en un dentífrico, o se puede agregar en una cantidad terapéuticamente efectiva a una composición que puede incluir agua, aglutinantes, abrasivos, agentes saborizantes, agentes espumantes y humectantes. Alternativamente, las composiciones se pueden configurar en una forma de comprimido o de solución, que se puede colocar debajo de la lengua o se puede disolver de otra manera en la boca.

2. Suministro inyectable.

15 En ciertas circunstancias, será deseable administrar las composiciones farmacéuticas dadas a conocer en la presente memoria, parenteralmente, intravenosamente, intramuscularmente, o inclusive intraperitonealmente, como se describe en la Patente de los EE.UU. n.º: 5.543.158; en la Patente de los EE.UU. n.º: 5.641.515 y en la Patente de los EE.UU. n.º: 5.399.363 (cada una incorporada específicamente a la presente memoria por referencia en su totalidad). Las soluciones de los compuestos activos como base libre o como sales farmacológicamente aceptables, se pueden preparar en agua adecuadamente mezclada con un tensioactivo, tal como hidroxipropilcelulosa. Las dispersiones también se pueden preparar en glicerol, polietilenglicoles líquidos y mezclas de los mismos y en aceites. Según condiciones ordinarias de almacenamiento y uso, estas preparaciones contienen un conservador para prevenir el crecimiento de microorganismos.

25 Las formas farmacéuticas adecuadas para uso inyectable incluyen soluciones o dispersiones acuosas estériles y polvos estériles para la preparación extemporánea de soluciones o dispersiones inyectables estériles (Patente de los EE.UU. n.º: 5.466.468, incorporada específicamente a la presente memoria por referencia en su totalidad). En todos los casos, la forma debe ser estéril y debe ser fluida hasta el grado en que exista una fácil posibilidad de pasarse por jeringa. Debe ser estable según las condiciones de fabricación y almacenamiento y se debe conservar contra la acción contaminante de los microorganismos, tales como bacterias y hongos. El vehículo puede ser un solvente o un medio de dispersión que contenga, por ejemplo, agua, etanol, poliol (por ejemplo, glicerol, propilenglicol y polietilenglicol líquido y similares), mezclas adecuadas de los mismos, y/o aceites vegetales. Se puede mantener la fluidez apropiada, por ejemplo, mediante el uso de un recubrimiento, tal como lecitina, mediante el mantenimiento del tamaño de partículas requerido en el caso de la dispersión y mediante el uso de tensioactivos. La prevención de la acción de los microorganismos se puede facilitar mediante diferentes agentes antibacterianos y antifúngicos, por ejemplo parabenos, clorbutanol, fenol, ácido sórbico, timerosal y similares. En muchos casos, será preferible incluir agentes isotónicos, por ejemplo azúcares o cloruro de sodio. Se puede provocar una absorción prolongada de las composiciones inyectables mediante el uso en las composiciones de agentes que demoren la absorción, por ejemplo, monoestearato de aluminio y gelatina.

40 Para la administración parenteral en una solución acuosa, por ejemplo, la solución debe regularse adecuadamente, si es necesario y el diluyente líquido primeramente se hace isotónico con suficiente suero o glucosa. Estas soluciones acuosas particulares son en especial adecuadas para la administración intravenosa, intramuscular, subcutánea, e intraperitoneal. En relación con esto, un medio acuoso estéril que se puede emplear será conocido por los expertos en este campo, a la luz de la presente divulgación. Por ejemplo, una dosificación se puede disolver en 1 mililitro de una solución isotónica de NaCl y se agrega a 1.000 mililitros de fluido de hipodermoclasia, o se inyecta en el sitio propuesto de infusión (véase, por ejemplo, *Remington's Pharmaceutical Sciences*, 15ª Edición, páginas 1035-1038 y 1570-1580). Necesariamente se presentará alguna variación en la dosificación, dependiendo de la condición del sujeto se esté tratando. La persona responsable de la administración, en cualquier caso, determinará la dosis apropiada para el sujeto individual. Más aún, para la administración a seres humanos, las preparaciones deben satisfacer las normas de esterilidad, pirogenicidad y de seguridad y pureza generales requeridas por la Oficina de Normas Biológicas de la FDA.

55 Las soluciones inyectables estériles se preparan mediante la incorporación de los compuestos activos en la cantidad requerida en el solvente apropiado con otros diferentes ingredientes enumerados anteriormente, según se requieran, seguida por esterilización filtrada. En términos generales, las dispersiones se preparan mediante la incorporación de los diferentes ingredientes activos esterilizados en un vehículo estéril que contenga el medio de dispersión básico y los otros ingredientes requeridos a partir de los enumerados anteriormente. En el caso de los polvos estériles para la preparación de soluciones inyectables estériles, los procedimientos de preparación preferidos son las técnicas de secado al vacío y secado por congelación, que proporcionan un polvo del ingrediente activo más cualquier ingrediente deseado adicional a partir de una solución previamente filtrada estéril de los mismos.

Las composiciones dadas a conocer en la presente memoria se pueden formular en una forma neutra o de sal. Las sales farmacéuticamente aceptables incluyen las sales de adición de ácido (formadas con los grupos amino libres de la proteína) y las que se forman con ácidos inorgánicos, tales como, por ejemplo, ácidos clorhídrico o fosfórico, o con ácidos orgánicos, tales como ácido acético, oxálico, tartárico, mandélico y similares. Las sales formadas con los grupos carboxilo libres también se pueden derivar a partir de bases inorgánicas, tales como, por ejemplo, hidróxidos de sodio, potasio, amonio, calcio, o férrico y las bases orgánicas tales como isopropil-amina, trimetil-amina, histidina, procaína y similares. Después de la formulación, las soluciones se administrarán de una manera compatible con la formulación de dosificación y en una cantidad tal que sea terapéuticamente efectiva. Las formulaciones se administran fácilmente en una variedad de formas de dosificación, tales como soluciones inyectables, cápsulas de liberación de fármaco y similares.

Como se utiliza en la presente memoria, "vehículo" incluye cualquiera y todos los solventes, medios de dispersión, vehículos, recubrimientos, diluyentes, agentes antibacterianos y antifúngicos, agentes isotónicos y de demora de absorción, reguladores del pH, soluciones portadoras, suspensiones, coloides y similares. El uso de estos medios y agentes para las sustancias activas farmacéuticas es bien conocido en la materia. Excepto hasta donde cualquier medio o agente convencional sea incompatible con el ingrediente activo, se contempla su uso en las composiciones terapéuticas. También se pueden incorporar ingredientes activos complementarios en las composiciones.

La frase "farmacéuticamente aceptable" se refiere a las entidades moleculares y composiciones que no produzcan una reacción alérgica o similar indeseada cuando se administren a un ser humano. La preparación de una composición acuosa que contenga una proteína como un ingrediente activo es bien entendida en la técnica. Típicamente, estas composiciones se preparan como inyectables, ya sea como soluciones o suspensiones líquidas; también se pueden preparar formas sólidas adecuadas para solución en, o suspensión en, un líquido antes de la inyección. La preparación también se puede emulsionar.

3. Suministro nasal y bucal.

En ciertas realizaciones, las composiciones farmacéuticas se pueden administrar mediante aerosoles intranasales, aerosoles bucales, inhalación, y/u otros vehículos de suministro en aerosol. Los procedimientos para administrar genes, ácidos nucleicos y composiciones peptídicas directamente a los pulmones, por ejemplo por medio de rocíos de aerosol nasal y bucal, ya se han descrito, por ejemplo, en la Patente de los EE.UU. n.º: 5.756.353 y en la Patente de los EE.UU. n.º: 5.804.212 (cada una específicamente incorporada a la presente memoria por referencia en su totalidad). De la misma manera, la administración de fármacos utilizando resinas de micropartículas intranasales (Takenaga y colaboradores, 1998) y compuestos de lisofosfatidil-glicerol (Patente de los EE.UU. n.º: 5.725.871, específicamente incorporada a la presente memoria por referencia en su totalidad), también es bien conocido en la técnica farmacéutica. De la misma manera, la administración de fármacos transmucosos en la forma de una matriz de soporte de politetra-fluoro-etileno se describe en la Patente de los EE.UU. n.º: 5.780.045 (específicamente incorporada a la presente memoria por referencia en su totalidad).

4. Suministro mediado por liposomas, nanocápsulas y micropartículas.

En ciertas realizaciones, los autores de la invención contemplan el uso de liposomas, nanocápsulas, micropartículas, microesferas, partículas de lípido, vesículas y similares, para la introducción de las composiciones de la presente memoria en células huésped adecuadas. En particular, las composiciones de la presente invención se pueden formular para administrarse encapsuladas en una partícula de lípido, un liposoma, una vesícula, una nanoesfera, o bien en una nanopartícula, o similares.

Estas formulaciones se pueden preferir para la introducción de formulaciones farmacéuticamente aceptables de los ácidos nucleicos o las construcciones que se dan a conocer en la presente memoria. La formación y uso de liposomas son conocidos en general por los expertos en este campo (véase, por ejemplo, Couvreur y colaboradores, 1977; Couvreur y colaboradores, 1988; Lasic, 1998; que describe el uso de liposomas y nanocápsulas en la terapia con antibióticos dirigidos para infecciones y enfermedades bacterianas intracelulares). Recientemente, se desarrollaron liposomas con una mejor estabilidad en suero y mejores tiempos medios de circulación (Gabizon y Papahadjopoulos, 1988; Allen y Choun, 1987; Patente de los EE.UU. n.º: 5.741.516, específicamente incorporada a la presente memoria por referencia en su totalidad). Además, se han revisado diferentes procedimientos de preparaciones de liposomas y tipo liposomas como vehículos de fármacos potenciales (Takakura, 1998; Chandran y colaboradores, 1997; Margalit, 1995; Patente de los EE.UU. n.º: 5.567.434; Patente de los EE.UU. n.º: 5.552.157; Patente de los EE.UU. n.º: 5.565.213; Patente de los EE.UU. n.º: 5.738.868 y Patente de los EE.UU. n.º: 5.795.587, cada una específicamente incorporada a la presente memoria por referencia en su totalidad).

Los liposomas se han utilizado con éxito con un número de tipos de células que normalmente son resistentes a la transfección mediante otros procedimientos, incluyendo suspensiones de células-T, cultivos de hepatocitos primarios y células PC 12 (Renneisen y colaboradores, 1990; Muller y colaboradores, 1990). Además, los liposomas están libres de las limitaciones de la longitud del ADN que son típicas de los sistemas de suministro basados en virus. Los liposomas se han utilizado efectivamente para introducir genes, fármacos (Heath y Martin, 1986; Heath y

colaboradores, 1986; Balazsovits y colaboradores, 1989; Fresta y Puglisi, 1996), agentes radioterapéuticos (Pikul y colaboradores, 1987), enzimas (Imaizumi y colaboradores, 1990a; Imaizumi y colaboradores, 1990b), virus (Faller & Baltimore, 1984), factores de transcripción y efectores aloestéricos (Nicolau y Gersonde, 1979), en una variedad de líneas celulares cultivadas y animales. Además, se han llevado a cabo varios estudios clínicos de éxito que examinan la efectividad de la administración de fármacos mediado por liposomas (López-Berestein y colaboradores, 1985a; 1985b; Coune, 1988; Sculier y colaboradores, 1988). Adicionalmente, varios estudios sugieren que el uso de liposomas no está asociado con las respuestas autoinmunes, la toxicidad, o la localización gonadal después de la administración sistémica (Mori y Fukatsu, 1992).

Los liposomas se forman a partir de fosfolípidos que se dispersan en un medio acuoso y forman espontáneamente vesículas de bicapa concéntricas multilamelares (también denominadas como vesículas multilamelares (MLV)). Las vesículas multilamelares generalmente tienen diámetros de 25 nanómetros a 4 micrómetros. La sonicación de las vesículas multilamelares da como resultado la formación de pequeñas vesículas unilamelares (SUV) con diámetros en el intervalo de 200 a 500 Å, que contienen una solución acuosa en el núcleo.

Los liposomas tienen un parecido con las membranas celulares y se contemplan para utilizarse en relación con la presente invención como vehículos para las composiciones peptídicas. Son ampliamente adecuados, debido a que se pueden atrapar tanto las sustancias tanto solubles en agua como solubles en lípido, es decir, en los espacios acuosos y dentro la bicapa misma, respectivamente. Es posible que se puedan emplear inclusive los liposomas que lleven fármacos para la administración específico del sitio de los agentes activos, mediante la modificación selectiva de la formulación liposomal.

Además de las enseñanzas de Couvreur y colaboradores (1977; 1988), se puede utilizar la siguiente información para generar formulaciones liposomales. Los fosfolípidos pueden formar una variedad de estructuras diferentes de los liposomas cuando se dispersan en agua, dependiendo de la proporción molar del lípido al agua. En bajas proporciones, el liposoma es la estructura preferida. Las características físicas de los liposomas dependen del pH, la concentración iónica y la presencia de cationes divalentes. Los liposomas pueden mostrar una baja permeabilidad a las sustancias iónicas y polares, pero a temperaturas elevadas, sufren una transición de fases que altera notoriamente su permeabilidad. La transición de fases involucra un cambio desde una estructura ordenada estrechamente empacada, conocida como el estado de gel, hasta una estructura menos ordenada, sueltamente empacada, conocida como el estado fluido. Esto ocurre a una temperatura de transición de fases característica y da como resultado un aumento en la permeabilidad a los iones, a los azúcares y a los fármacos.

Además de la temperatura, la exposición a las proteínas puede alterar la permeabilidad de los liposomas. Ciertas proteínas solubles, tales como el citocromo C, se enlazan con, deforman y penetran en la bicapa, causando de esta manera cambios en la permeabilidad. El colesterol inhibe esta penetración de proteínas, aparentemente por el empaque de los fosfolípidos que está más apretado. Se contempla que las formaciones de liposomas más útiles para la administración de antibióticos e inhibidores, contendrán colesterol.

La capacidad para atrapar solutos varía entre diferentes tipos de liposomas. Por ejemplo, las vesículas multilamelares son moderadamente eficientes para atrapar solutos, pero las vesículas unilamelares pequeñas son extremadamente ineficientes. Las vesículas unilamelares pequeñas ofrecen la ventaja de la homogeneidad y reproducibilidad en la distribución de tamaños; sin embargo, las vesículas unilamelares grandes ofrecen un compromiso entre el tamaño y la eficiencia de atrape (LUV). Estas se preparan mediante evaporación de éter y son de tres a cuatro veces más eficientes para atrapar el soluto que las vesículas multilamelares.

Además de las características de los liposomas, un determinante importante en el atrape de compuestos es el de las propiedades fisicoquímicas del compuesto mismo. Los compuestos polares se atrapan en los espacios acuosos y los compuestos no polares se enlazan con la bicapa de lípido de la vesícula. Los compuestos polares se liberan a través de la permeación o cuando se rompe la bicapa, pero los compuestos no polares permanecen afiliados con la bicapa, a menos que sea alterada por la temperatura o la exposición a las lipoproteínas. Ambos tipos muestran máximos índices de eflujo a la temperatura de transición de fases.

Los liposomas interactúan con las células por medio de cuatro mecanismos diferentes: endocitosis por parte de las células fagocíticas del sistema reticuloendotelial, tales como macrófagos y neutrófilos; adsorción a la superficie celular, ya sea mediante las fuerzas hidrofóbicas o electrostáticas débiles no específicas, o bien mediante interacciones específicas con los componentes de superficie celular; fusión con la membrana celular de plasma mediante la inserción de la bicapa de lípido del liposoma en la membrana de plasma, con la liberación simultánea del contenido liposomal hacia el citoplasma; y mediante transferencia de los lípidos liposomales hacia las membranas celulares o subcelulares, o *viceversa*, sin asociación alguna del contenido del liposoma. Con frecuencia es difícil determinar cuál mecanismo es operativo y más de uno pueden operar al mismo tiempo.

El objetivo y la disposición de los liposomas intravenosamente inyectados dependen de sus propiedades físicas, tales como su tamaño, fluidez y carga superficial. Pueden persistir en los tejidos durante horas o días, dependiendo de su composición y las semividas en la sangre están en el intervalo desde minutos hasta varias horas. Los

liposomas más grandes, tales como las vesículas multilamelares y las vesículas unilamelares grandes, son rápidos absorbidos por las células fagocíticas del sistema reticuloendotelial, pero la fisiología del sistema circulatorio restringe la salida de estas especies grandes en la mayoría de los sitios. Pueden salir solamente en los lugares en donde existan grandes aberturas o poros en el endotelio capilar, tal como los senos del hígado o del bazo. Por consiguiente, estos órganos son el sitio predominante de absorción. Por otra parte, las vesículas unilamelares pequeñas muestran una distribución de tejido más amplia, pero todavía son altamente secuestradas en el hígado y el bazo. En general, este comportamiento *in vivo* limita la dirección potencial de los liposomas hacia solamente los órganos y tejidos accesibles para su gran tamaño. Estos incluyen la sangre, el hígado, el bazo, la médula ósea y los órganos linfoides.

La dirección en general no es una limitación en términos de la presente memoria. Sin embargo, si se desea una dirección específica, existen procedimientos disponibles para que esto se lleve a cabo. Se pueden utilizar anticuerpos para enlazarse con la superficie del liposoma y para dirigir el anticuerpo y su contenido de fármaco hacia los receptores antigénicos específicos localizados sobre una superficie de tipo celular particular. También se pueden utilizar determinantes de carbohidratos (componentes de superficie celular de glucoproteína o glucolípido que tiene un papel en el reconocimiento, interacción y adhesión de célula-célula) como sitios de reconocimiento, debido a que tienen el potencial para dirigir los liposomas hacia tipos particulares de células. En su mayor parte, se contempla que se utilizaría inyección intravenosa de las preparaciones liposomales, pero también son concebibles otras vías de administración.

De una manera alternativa, la invención proporciona formulaciones de nanocápsulas farmacéuticamente aceptables de las composiciones de la presente memoria. Las nanocápsulas pueden atrapar en general los compuestos de una manera estable y reproducible (Henry-Michelland y colaboradores, 1987; Quintanar-Guerrero y colaboradores, 1998; Douglas y colaboradores, 1987). Para evitar los efectos secundarios debidos a la sobrecarga polimérica intracelular, estas partículas ultrafinas (de un tamaño de alrededor de 0.1 micrómetros) deben diseñarse utilizando polímeros capaces de degradarse *in vivo*. Las nanopartículas de poli-alquil-cianoacrilato biodegradables que satisfacen estos requerimientos son contempladas para utilizarse en la presente memoria. Estas partículas se pueden hacer fácilmente como se describe (Couvreur y colaboradores, 1980; 1988; zur Muhlen y colaboradores, 1998; Zambaux y colaboradores, 1998; Pinto-Alphandry y colaboradores, 1995 y Patente de los EE.UU. n.º: 5.145.684, específicamente incorporados a la presente memoria por referencia en su totalidad).

Vacunas

En ciertas realizaciones preferidas de la presente invención, se proporcionan vacunas. Las vacunas generalmente comprenderán una o más composiciones farmacéuticas, tales como las discutidas anteriormente, en combinación con un inmunoestimulante. Un inmunoestimulante puede ser cualquier sustancia que mejore o potencie una respuesta inmune (mediada por anticuerpo y/o por células) a un antígeno exógeno. Los ejemplos de los inmunoestimulantes incluyen adyuvantes, microesferas biodegradables (por ejemplo, galactida poliláctica) y liposomas (en los que se incorpora el compuesto; véase, por ejemplo, Fullerton, Patente de los EE.UU. n.º: 4.235.877). La preparación de vacuna se describe en general, por ejemplo, por Powell y Newman, editores, *Vaccine Design* (el planteamiento de la subunidad y adyuvante) (1995). Las composiciones farmacéuticas y vacunas también pueden contener otros compuestos, que pueden ser biológicamente activos o inactivos. Por ejemplo, puede haber una o más porciones inmunógenas de otros antígenos tumorales presentes, ya sea incorporados en un polipéptido de fusión, o bien como un compuesto separado, dentro de la composición o vacuna.

Las vacunas ilustrativas pueden contener ADN que codifique uno o más de los polipéptidos, como se describe anteriormente, de tal manera que el polipéptido se genere *in situ*. Como se observó anteriormente, el ADN puede estar presente dentro de cualquiera de una variedad de sistemas de suministro conocidos por los expertos en la materia, incluyendo sistemas de expresión de ácido nucleico, bacterias y sistemas de expresión viral. En este campo se conocen numerosas técnicas de suministro de genes, tales como las descritas por Rolland, *Crit. Rev. Therap. Drug Carrier Systems* 15:143-198 (1998) y las referencias citadas en el mismo. Los sistemas de expresión de ácido nucleico apropiados contienen las secuencias de ADN necesarias para la expresión en el paciente (tal como una señal promotora y terminadora adecuada). Los sistemas de suministro bacteriano involucran la administración de una célula huésped bacteriana (por ejemplo, una cepa de *Mycobacterium*, *Bacillus* o *Lactobacillus*, incluyendo *Bacillus Calmette-Guerin* o *Lactococcus lactis*), que exprese una porción inmunógena del polipéptido sobre su superficie celular, o que secrete este epitopo (véase, por ejemplo, Ferreira y colaboradores, *An Acad. Bras. Cienc.* (2005) 77:113-124; y Raha y colaboradores, *Appl. Microbiol. Biotechnol.* (2005) PubMedID 15635459). En una realización preferida, el ADN se puede introducir utilizando un sistema de expresión viral (por ejemplo, vacuna u otro poxvirus, retrovirus, o adenovirus), que puede involucrar el uso de un virus no patogénico (defectuoso), competente en réplica. Los sistemas adecuados se dan a conocer, por ejemplo, por Fisher-Hoch y colaboradores, *Proc. Natl. Acad. Sci. EUA* 86:317-321 (1989); Flexner y colaboradores, *Ann. N. Y. Acad. Sci.* 569:86-103 (1989); Flexner y colaboradores, *Vaccine* 8:17-21 (1990); Patentes de los EE.UU. N.ºs 4.603.112, 4.769.330 y 5.017.487; documento WO 89/01973; Patente de los EE.UU. n.º: 4.777.127; Patente Británica n.º: GB 2.200.651; Patente Europea n.º: EP 0.345.242; documento WO 91/02805; Berkner, *Biotechniques* 6:616-627 (1988); Rosenfeld y colaboradores, *Science* 252:431-434 (1991); Kolls y colaboradores, *Proc. Natl. Acad. Sci. EUA* 91:215-219 (1994); Kass-Eisler y

colaboradores, *Proc. Natl. Acad. Sci. EUA* 90:11498-11502 (1993); Guzman y colaboradores, *Circulation* 88:2838-2848 (1993); y Guzman y colaboradores, *Cir. Res.* 73:1202-1207 (1993). Las técnicas para incorporar el ADN en estos sistemas de expresión son bien conocidas por los expertos en este campo. El ADN también puede estar "desnudo", como se describe, por ejemplo, por Ulmer y colaboradores, *Science* 259:1745-1749 (1993) y como es revisado por Cohen, *Science* 259:1691-1692 (1993). La absorción del ADN desnudo se puede aumentar recubriendo el ADN sobre perlas biodegradables, que se transportan eficientemente hacia dentro de las células. Será aparente que una vacuna puede comprender tanto un componente de polinucleótido como de polipéptido. Estas vacunas pueden proporcionar una mejor respuesta inmune.

Será aparente que una vacuna puede contener sales farmacéuticamente aceptables de los polinucleótidos y polipéptidos proporcionados en la presente memoria. Estas sales se pueden preparar 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).

Aunque se puede emplear cualquier vehículo adecuado conocido por los expertos en la técnica, en las composiciones de vacuna, el tipo de vehículo variará dependiendo del modo de administración. Las composiciones de la presente invención se pueden formular para cualquier modo de administración apropiado, incluyendo, por ejemplo, administración tópica, oral, nasal, intravenosa, intracraneal, intraperitoneal, subcutánea, o intramuscular. Para la administración parenteral, tal como inyección subcutánea, el vehículo de preferencia comprende agua, suero, alcohol, una grasa, una cera, o un regulador. Para la administración oral, se puede emplear cualquiera de los vehículos anteriores, o un vehículo sólido, tal como manitol, lactosa, almidón, estearato de magnesio, sacarina de sodio, talco, celulosa, glucosa, sacarosa y carbonato de magnesio. También se pueden emplear microesferas biodegradables (por ejemplo, polilactato-poliglicolato) como vehículos para las composiciones farmacéuticas de esta invención. Las microesferas biodegradables adecuadas se dan a conocer, por ejemplo, en las Patentes de los EE.UU. n.ºs: 4.897.268; 5.075.109; 5.928.647; 5.811.128; 5.820.883; 5.853.763; 5.814.344 y 5.942.252. También se puede emplear un vehículo que comprenda los complejos de partículas-proteína descritos en la Patente de los EE.UU. n.º: 5.928.647, que son capaces de inducir una respuesta de linfocitos-T citotóxica restringida a la clase I en un huésped.

Estas composiciones también pueden comprender reguladores del pH (por ejemplo, suero regulado neutro o suero regulado con fosfato), carbohidratos (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 hagan a la formulación isotónica, hipotónica, o débilmente hipertónica con la sangre de un receptor, agentes de suspensión, agentes espesantes, y/o conservadores. De una manera alternativa, las composiciones de la presente invención se pueden formular como un liofilizado. Los compuestos también se pueden encapsular dentro de liposomas empleando la tecnología bien conocida.

Se puede emplear cualquiera de una variedad de inmunoestimulantes en las vacunas de esta invención. Por ejemplo, se puede incluir un adyuvante. La mayoría de los adyuvantes contienen una sustancia diseñada para proteger al antígeno del catabolismo rápido, tal como hidróxido de aluminio o aceite mineral y un estimulante de las respuestas inmunes, tal como lípido A, las especies de *Bordetella pertussis* o *Mycobacterium*, o proteínas derivadas a partir de *M. tuberculosis*. Por ejemplo, se puede utilizar *M. vaccae* ("pVac") desglucolipidada, deslipidada. Los adyuvantes adecuados están comercialmente disponibles, por ejemplo, como Adyuvante Incompleto y Adyuvante Completo de Freund (Difco Laboratories, Detroit, MI); Adyuvante Merck 65 (Merck and Company, Inc., Rahway, NJ); AS01B, AS02A, AS15, AS-2 y sus derivados (GlaxoSmithKline, Filadelfia, PA); CWS, TDM, Leif, sales de aluminio, tales como gel de hidróxido de aluminio (alumbre), o fosfato de aluminio; sales de calcio, hierro, o zinc; una suspensión insoluble de tirosina acilada; azúcares acilados; polisacáridos catiónicamente o aniónicamente derivados; polifosfacenos; microesferas biodegradables; monofosforil-lípido A y quil A. También se pueden utilizar como adyuvantes las citocinas, tales como GM-CSF o interleucina-2, -7, o -12.

Dentro de las vacunas proporcionadas en la presente memoria, la composición del adyuvante de preferencia se diseña para introducir una respuesta inmune predominantemente del tipo Th1. Los altos niveles de citocinas tipo Th1 (por ejemplo, IFN- γ , TNF α , IL-2, e IL-12) tienden a favorecer la inducción de las respuestas inmunes mediadas por células a un antígeno administrado. En contraste, los altos niveles de citocinas tipo Th2 (por ejemplo, IL-4, IL-5, IL-6, e IL-10) tienden a favorecer la inducción de las respuestas inmunes humorales. A continuación de la aplicación de una vacuna como se dispone en la presente memoria, un paciente soportará una respuesta inmune que incluya respuestas tipo Th1 y Th2. Dentro de una realización preferida, en donde una respuesta es predominantemente del tipo Th1, el nivel de citocinas tipo Th1 aumentará hasta un mayor grado que el nivel de citocinas tipo Th2. Los niveles de estas citocinas se pueden evaluar fácilmente empleando ensayos convencionales. Para una revisión de las familias de citocinas, véase Janeway y colaboradores, *Immunobiology*, 5ª Edición, 2001.

Los adyuvantes preferidos para utilizarse para provocar una respuesta predominantemente tipo Th1 incluyen, por ejemplo, una combinación de fosforil-lípido A, de preferencia monofosforil-lípido A 3-O-desacilado (3D-MPL),

opcionalmente con una sal de aluminio (véase, por ejemplo, Ribi y colaboradores, 1986, *Immunology and Immunopharmacology of Bacterial Endotoxins*, Plenum Publ. Corp., NY, páginas 407-419; Patentes Británicas n.^{os} GB 2122204B y GB 2220211; y Patente de los EE.UU. n.º: US 4.912.094). Una forma preferida de 3D-MPL es la forma de una emulsión que tiene un tamaño de partículas pequeño menor de 0.2 milímetros de diámetro y su procedimiento de fabricación se da a conocer en el documento WO 94/21292. Las formulaciones acuosas que comprenden monofosforil-lípido A y un tensioactivo se han descrito en el documento WO 98/43670. Los adyuvantes preferidos ejemplificados incluyen AS01B (MPL y QS21 en una formulación de liposomas), 3D-MPL y QS21 en una formulación de liposomas, AS02A (MPL y QS21 y una emulsión de aceite en agua), 3D-MPL y QS21 y una emulsión de aceite en agua y AS15, disponible en GlaxoSmithKline. Los adyuvantes de MPL están disponibles en GlaxoSmithKline, Seattle, WA (ver las Patentes de los EE.UU. n.^{os} 4.436.727; 4.877.611; 4.866.034 y 4.912.094).

Los oligonucleótidos que contienen CpG (en donde el dinucleótido CpG no está metilado) también inducen una respuesta predominantemente Th1. CpG es una abreviatura para los motivos de los dinucleótidos de citosina-guanosina presentes en el ADN. Estos oligonucleótidos son bien conocidos y se describen, por ejemplo, en las publicaciones internacionales WO 96/02555 y WO 99/33488 y en las Patentes de los EE.UU. n.^{os} 6.008.200 y 5.856.462. También se describen secuencias de ADN inmunoestimulantes, por ejemplo, por Sato y colaboradores, *Science* 273:352 (1996). CpG, cuando se formula en vacunas, se administra en general en solución libre junto con el antígeno libre (documento WO96/02555; McCluskie y Davis, *supra*), o covalentemente conjugado con un antígeno (documento WO 98/16247), o se formula con un vehículo, tal como hidróxido de aluminio ((antígeno superficial de hepatitis), Davis y colaboradores, *supra*; Brazolot-Millan y colaboradores, *Proc. Natl. Acad. Sci. EUA*, 1998, 95(26), 15553-8). CpG se conoce en la técnica como un adyuvante que se puede administrar tanto mediante la vía sistémica como mucosa (documento WO 96/02555, Patente Europea n.º: EP 468520, Davis y colaboradores, *J. Immunol.*, 1998, 160(2):870-876; McCluskie y Davis, *J. Immunol.*, 1998, 161(9):4463-6).

Otro adyuvante preferido es una saponina o miméticos o derivados de saponina, de preferencia QS21 (Aquila Biopharmaceuticals Inc., Framingham, MA), que se puede utilizar sola o en combinación con otros adyuvantes. Por ejemplo, un sistema mejorado involucra la combinación de un monofosforil-lípido A y un derivado de saponina, tal como la combinación de QS21 y 3D-MPL, como se describe en el documento WO 94/00153, o una composición menos reactogénica, en donde el QS21 se apaga con colesterol, como se describe en el documento WO 96/33739. Otras formulaciones preferidas comprenden una emulsión de aceite en agua y tocoferol. Una formulación de adyuvante particularmente potente que involucra QS21, 3D-MPL y tocoferol en una emulsión de aceite en agua, se describe en el documento WO 95/17210. Los adyuvantes de saponina adicionales útiles en la presente invención incluyen QS7 (descrito en las publicaciones internacionales WO 96/33739 y WO 96/11711) y QS17 (descrito en la Patente de los EE.UU. n.º: 5.057.540 y en la Patente Europea n.º: EP 0.362.279 B1).

Otros adyuvantes preferidos incluyen Montanide ISA 720 (Seppic, Francia), SAF (Chiron, California, EE.UU.), ISCOMS (CSL), MF-59 (Chiron), la serie de adyuvantes SBAS (por ejemplo, SBAS-2, AS2', AS2'', SBAS-4, o SBAS6, disponibles en GlaxoSmithKline, Rixensart, Bélgica), Detox (Corixa, Hamilton, MT), RC-529 (Corixa, Hamilton, MT) y otros 4-fosfatos de amino-alkil-glucosaminida (AGP), tales como los descritos en las Solicitudes de Patente de los EE.UU. en tramitación con Números de Serie 08/853.826 y 09/074.720, cuyas divulgaciones se incorporan a la presente memoria por referencia en su totalidad.

Otros adyuvantes de ejemplo incluyen MPL sintético y los adyuvantes basados en la subunidad B de toxina Shiga (véase el documento WO2005/112991).

Cualquier vacuna proporcionada en la presente memoria se puede preparar empleando procedimientos bien conocidos que den como resultado una combinación de antígeno, potenciador de la respuesta inmune y un vehículo o excipiente adecuado. Las composiciones descritas en la presente memoria se pueden administrar como parte de una formulación de liberación sostenida (es decir, una formulación tal como una cápsula, esponja, o gel (compuesto de polisacáridos, por ejemplo), que efectúe una liberación lenta del compuesto a continuación de la administración). Estas formulaciones se pueden preparar en general empleando la tecnología bien conocida (véase, por ejemplo, Coombes y colaboradores, *Vaccine* 14:1429-1438 (1996)) y se administran, por ejemplo, mediante implante oral, rectal, o subcutáneo, o mediante implante en el sitio objetivo deseado. Las formulaciones de liberación sostenida pueden contener un polipéptido, polinucleótido, o anticuerpo dispersado en una matriz de vehículo, y/o pueden estar contenidos dentro de un depósito rodeado por una membrana de control de velocidad.

Los vehículos para utilizarse dentro de estas formulaciones son biocompatibles y también pueden ser biodegradables; de preferencia la formulación proporciona un nivel relativamente constante de liberación del componente activo. Estos vehículos incluyen micropartículas de poli-(láctido-co-glicólido), poliacrilato, látex, almidón, celulosa, dextrano y similares. Otros vehículos de liberación retardada incluyen biovectores supramoleculares, que comprenden un núcleo hidrófilo no líquido (por ejemplo, un polisacárido u oligosacárido reticulado) y opcionalmente, una capa externa que comprende un compuesto anfifílico, tal como un fosfolípido (véase, por ejemplo, la Patente de los EE.UU. n.º: 5.151.254 y las solicitudes del TCP WO 94/20078, WO/94/23701 y WO 96/06638). La cantidad de compuesto activo contenida dentro de una formulación de liberación sostenida depende del sitio de implante, de la velocidad y duración esperada de liberación y de la naturaleza de la condición que se vaya a tratar o a prevenir.

Se puede emplear cualquiera de una variedad de vehículos de suministro dentro de las composiciones farmacéuticas y vacunas para facilitar la producción de una respuesta inmune específica del antígeno que se dirija a las células tumorales. Los vehículos de suministro incluyen células presentadoras de antígeno (APC), tales como células dendríticas, macrófagos, células-B, monocitos y otras células que se puedan diseñar para ser células presentadoras de antígeno eficientes. Estas células pueden, pero no necesitan, modificarse genéticamente para aumentar la capacidad para presentar el antígeno, para mejorar la activación y/o el mantenimiento de la respuesta de células-T, para tener efectos anti-tumorales por sí mismas, y/o para ser inmunológicamente compatibles con el receptor (es decir, haplotipo HLA emparejado). Las células presentadoras de antígeno se pueden aislar en general a partir de cualquiera de una variedad de fluidos biológicos y órganos, incluyendo tejidos tumorales y peritumorales y pueden ser células autólogas, alogeneicas, singeneicas, o xenogeneicas.

Ciertas realizaciones preferidas de la presente invención utilizan células dendríticas o progenitoras de las mismas como células presentadoras de antígeno. Las células dendríticas son células presentadoras de antígeno altamente potentes (Banchereau y Steinman, *Nature* 392:245-251 (1998)) y se ha demostrado que son efectivas como un adyuvante fisiológico para provocar una inmunidad antitumoral profiláctica o terapéutica (véase Timmerman y Levy, *Ann. Rev. Med.* 50:507-529 (1999)). En general, las células dendríticas se pueden identificar basándose en su forma típica (de estrella *in situ*, con procesos citoplásmicos marcados (dendritas) visibles *in vitro*), su capacidad para absorber, procesar y presentar antígenos con una alta eficiencia y su capacidad para activar las respuestas de células-T puras. Desde luego, las células dendríticas se pueden diseñar para expresar receptores de superficie celular específicos o ligandos que comúnmente no se encuentran en las células dendríticas *in vivo* o *ex vivo*. Como una alternativa a las células dendríticas, se pueden utilizar células dendríticas cargadas con antígeno en vesículas secretadas (denominadas como exosomas) dentro de una vacuna (véase Zitvogel y colaboradores, *Nature Med.* 4:594-600 (1998)).

Las células dendríticas y sus progenitoras se pueden obtener a partir de sangre periférica, médula ósea, células infiltrantes de tumor, células infiltrantes de tejidos peritumorales, nodos linfáticos, bazo, piel, sangre de cordón umbilical, o cualquier otro tejido o fluido adecuado. Por ejemplo, las células dendríticas se pueden diferenciar *ex vivo* mediante la adición de una combinación de citocinas, tales como GM-CSF, IL-4, IL-13, y/o TNF α a los cultivos de monocitos cosechados de sangre periférica. De una manera alternativa, las células positivas para CD34 cosechadas de sangre periférica, sangre de cordón umbilical, o médula ósea, se pueden diferenciar en células dendríticas mediante la adición al medio de cultivo de combinaciones de GM-CSF, IL-3, TNF α , ligando CD40, LPS, ligando fIt3, y/u otros compuestos que induzcan diferenciación, maduración y proliferación de células dendríticas.

Las células dendríticas convenientemente se categorizan como células "inmaduras" y "maduras", lo que permite tener una manera simple de discriminar entre dos fenotipos bien caracterizados. Sin embargo, esta nomenclatura no debe interpretarse para excluir todas las posibles etapas intermedias de diferenciación. Las células dendríticas inmaduras se caracterizan como células presentadoras de antígeno con una alta capacidad para absorber y procesar el antígeno, lo que se correlaciona con la alta expresión del receptor Fc γ y el receptor de manosa. El fenotipo maduro típicamente se caracteriza por una expresión más baja de estos marcadores, pero una alta expresión de las moléculas de superficie celular responsables de la activación de células-T, tales como MHC clase I y clase II, moléculas de adhesión (por ejemplo, CD54 y CD11) y moléculas co-estimulantes (por ejemplo, CD40, CD80, CD86 y 4-1BB).

Las APC se pueden transfectar en general con un polinucleótido que codifique una proteína (o una porción u otra variante de la misma), de tal manera que el polipéptido, o una porción inmunógena del mismo, se exprese sobre la superficie celular. Esta transfección puede tener lugar *ex vivo* y entonces se puede utilizar una composición o vacuna que comprenda estas células transfectadas para propósitos terapéuticos, como se describe en la presente memoria. De una manera alternativa, se puede administrar un vehículo de suministro de genes que dirija una célula dendrítica u otra célula presentadora de antígeno a un paciente, dando como resultado que se presente la transfección *in vivo*. La transfección *in vivo* y *ex vivo* de células dendríticas, por ejemplo, se puede llevar a cabo en general empleando cualesquiera procedimientos conocidos en la materia, tales como los descritos en el documento WO 97/24447, o el planteamiento de la pistola de genes descrito por Mahvi y colaboradores, *Immunology and Cell Biology* 75:456-460 (1997). La carga de antígeno de las células dendríticas se puede lograr mediante la incubación de las células dendríticas o de las células progenitoras con el polipéptido, ADN (desnudo o dentro de un vector de plásmido), o ARN; o con bacterias o virus recombinantes que expresen el antígeno (por ejemplo, vectores de vacuna, varicela de aves, adenovirus, o lentivirus). Antes de cargarse, el polipéptido se puede conjugar covalentemente con un componente inmunológico que proporcione ayuda de células-T (por ejemplo, una molécula portadora). De una manera alternativa, una célula dendrítica se puede impulsar con un componente inmunológico no conjugado, por separado, o en la presencia del polipéptido.

Las vacunas y composiciones farmacéuticas se pueden presentar en recipientes de dosis unitaria o de múltiples dosis, tales como ampollitas o frascos sellados. Estos recipientes de preferencia se sellan herméticamente para conservar la esterilidad de la formulación hasta su uso. En general, las formulaciones se pueden almacenar como suspensiones, soluciones, o emulsiones en vehículos oleosos o acuosos. De una manera alternativa, una vacuna o una composición farmacéutica se puede almacenar en una condición secada por congelación, que solamente

requiera de la adición de un vehículo líquido estéril inmediatamente antes de usarse.

Ejemplos

Los siguientes Ejemplos se proporcionan a manera de ilustración solamente y no a manera de limitación. Los expertos en este campo reconocerán fácilmente una variedad de parámetros no críticos que se podrían cambiar o modificar para proporcionar resultados esencialmente similares.

Ejemplo de referencia 1: Preparación de Mtb72f (sin marca His) (SEC ID N.º:6)

Construcción del Vector de Expresión de Mtb72f.

Mtb72f es una proteína de fusión formada a partir de dos proteínas de *M. tuberculosis*, Mtb32 y Mtb39. Mtb72f se construye fusionando Mtb39 y las porciones N- y C-terminales de Mtb32 como sigue: extremo C-terminal de Mtb32 - Mtb39 - extremo N-terminal de Mtb32. Específicamente, se generó la proteína Mtb72f mediante el enlace en secuencia en fila de los marcos de lectura abierta (ORF) que codifican el fragmento C-terminal de aproximadamente 14 kDa de Mtb32 (residuos 192-323; 132 aminoácidos) a la longitud completa del marco de lectura abierta de Mtb39, seguido por el término-C con la porción N-terminal de aproximadamente 20 kDa (residuos 1-195) de Mtb32. Esto se llevó a cabo utilizando oligonucleótidos específicos de la secuencia que contenían sitios de restricción únicos (EcoRI y EcoRV) y que están desprovistos de los codones de paro en los extremos C-terminales (en el caso de Mtb32-C y Mtb39) para desactivar la reacción en cadena de la polimerasa (PCR) del ADN genómico a partir de la cepa H37Rv de *M. tuberculosis*.

Los detalles del proceso son como sigue:

Primero, se clonó el ADN que codifica la porción C-terminal de Mtb32 (Mtb32C) a partir de H37Rv utilizando reacción en cadena de la polimerasa con los siguientes oligonucleótidos: 5' (5'-CAA-TTA-CAT-ATG-CAT-CAC-CAT-CAC-CAT-CAC-ACG-GCC-GCG-TCC-GAT-AAC-TTC-3') y 3' (5'-CTA-ATC-GAA-TCC-GGC-CGG-GGG-TCC-CTC-GGC-CAA-3'). El oligonucleótido 5' contenía un sitio de restricción NdeI (subrayado) que abarca el codón de inicio ATG. El oligonucleótido 3' contenía un sitio de restricción EcoRI (subrayado). Estos oligonucleótidos se utilizaron para amplificar Mtb32C, una porción de 396 nucleótidos de Mtb32 y el producto resultante se subclonó en los sitios NdeI y EcoRI de un vector de expresión. La digestión con EcoRI y EcoRV subsiguientemente linealizó el plásmido Mtb32C.

Para Mtb39, se utilizaron los siguientes oligonucleótidos para la amplificación con reacción en cadena de la polimerasa y clonación: 5' (5'-CTA-ATC-GAA-TTC-ATG-GTG-GAT-TTC-GGG-GCG-TTA-3') y 3' (5'-CTA-ATC-GAT-ATC-GCC-GGC-TGC-CGG-AGA-ATG-CGG-3'). El oligonucleótido 5' contenía un sitio de restricción EcoRI (subrayado), mientras que el oligonucleótido 3' contenía un sitio de restricción EcoRV (subrayado). La secuencia de codificación de longitud completa de Mtb39 se amplificó, se digirió y se sub-clonó dentro del marco corriente abajo de Mtb32c, utilizando el plásmido previamente digerido del primer paso.

Los oligonucleótidos 5' y 3' del fragmento N-terminal de Mtb32 se diseñaron como sigue: 5'-(5'-CTA-ATC-GAT-ATC-GCC-CCG-CCG-GCC-TTG-TCG-CAG-GAC-3') y 3' (5'-CTA-ATC-GAT-ATC-CTA-GGA-CGC-GGC-CGT-GTT-CAT-AC-3'). Ambos conjuntos de oligonucleótidos contenían un sitio de restricción EcoRV (subrayado), mientras que el oligonucleótido 3' también incluía un codón de paro (cursivas). Los oligonucleótidos se diseñaron para amplificar una porción de 585 pares de bases de Mtb32, que codifican el dominio N-terminal predicho de esta proteína. El producto de reacción en cadena de la polimerasa se sub-clonó en el plásmido de fusión Mtb32c-Mtb39. Entonces se verificó la orientación apropiada de los insertos y la ausencia de mutaciones mediante secuenciación del ADN.

Para la construcción final, utilizada para hacer el Banco de Células Maestras y el Banco de Células de Trabajo del Fabricante, se removió la marca de afinidad 6xHis mediante reacción en cadena de la polimerasa y se subclonó el marco de lectura abierta (ORF) para Mtb32f en pPDM, un vector de expresión derivado de pET. El marco de lectura abierta codifica para una poliproteína de aproximadamente 72 kDa (Mtb72f), con dominios organizados en el orden lineal: Mtb32C, Mtb39-Mtb32N. Entonces este ADN se transformó en la cepa HMS174 pLysS de *E. coli* y se utilizó para la prueba, el banco de células y la fabricación.

Producción de Sustancia de Fármaco a Granel de Mtb72f.

El proceso de fabricación para la producción de Mtb72f se resume como sigue:

- Fermentación seguida por cosecha de células mediante centrifugación, alteración de células (microfluidizador) y centrifugación para proporcionar un gránulo de cuerpos de inclusión;
- Purificación del gránulo de cuerpos de inclusión mediante extracción en urea 8M, seguida por Cromatografía de Flujo Rápido de Q Sepharose (QFF), Cromatografía de Hidroxiapatita de Cerámica (CHT), diafiltración y filtración

esterilizante, para proporcionar la sustancia de fármaco a granel purificada.

Fermentación

Las fermentaciones se llevan a cabo en un volumen de procesamiento de 10 litros. El fermentador se inocula con 300 mililitros de un cultivo en matraz de agitación de las células de siembra de procesamiento cultivadas a 37°C durante la noche. Tanto el inóculo como la fermentación utilizan un medio semi-definido con glicerol derivado de planta como la fuente de carbono primaria. La composición del medio se muestra en la tabla que se encuentra más adelante. Todos los componentes del medio se esterilizan mediante calentamiento a 121°C durante 20 minutos, o mediante filtración esterilizante. Durante la fermentación, el fermentador se mantiene a una temperatura de 37°C. Se dispersa aire a una velocidad de 5 litros estándares por minuto (SLPM). El pH del medio se mantiene en 7.0 mediante la adición automática de ácido (H₂SO₄) o de base (NaOH). El fermentador se programa para controlar el oxígeno disuelto en el 30 % mediante el ajuste automático de la agitación, mientras que se mantiene una agitación mínima de 200 revoluciones por minuto. El control de la espuma dentro del fermentador se logra mediante la adición automática de anti-espumante de silicona SAG-471 al 1,05 % (Witco Corp.). Cuando la densidad celular alcanza una densidad óptica (600 nanómetros) de aproximadamente 3,5, se agrega 3,5-isopropil-beta-D-tiogalactopiranosida (IPTG) al fermentador en una concentración de 1,0 mM. La IPTG induce la expresión del gen recombinante que codifica la proteína Mtb72f. A las 3,0 horas después de la inducción, el fermentador se enfría y las células se cosechan mediante centrifugación en botellas centrífugas de 1 litro.

Composición del Medio de Fermentación

Material	Concentración
Extracto de levadura	15 gramos/litro
Glicerol	30 gramos/litro
Sulfato de magnesio, heptahidrato (MgSO ₄ -)	0,5 gramos/litro
Fosfato de potasio, monobásico (KH ₂ PO ₄)	2,4 gramos/litro
Fosfato de sodio, dibásico (Na ₂ HPO ₄)	3,2 gramos/litro
Cloruro de amonio (NH ₄ Cl)	1,0 gramos/litro
Cloruro de sodio (NaCl)	0,5 gramos/litro
Sulfato de canamicina	30 miligramos/litro
Cloranfenicol	34 miligramos/litro
Antiespumante de silicona SAG-471 (Witco Corp.)	0,0005% (v/v) (no incluido)

20 Aislamiento de Cuerpos de Inclusión

Los gránulos se vuelven a suspender y se reservan en 2,3 litros de regulador de lisis (NaCl 50 mM, Tris 10 mM, pH de 8,0) y se utiliza un M-110Y Microfluidizer® para alterar las células. Las células se pasan a través del Microfluidizador cinco veces a una presión de 75,84 ± 6,89 MegaPascales (11.000 ± 1.000 psi). La suspensión se centrifuga a 8.000 x g en botellas de 500 mililitros. Según estas condiciones, los cuerpos de inclusión (IB) que contienen la proteína Mtb72f se forman en gránulos, mientras que la mayor parte del desecho celular permanece en el sobrenadante. Los gránulos de los cuerpos de inclusión se vuelven a suspender en regulador de lavado (urea 2 M, NaCl 50 mM, Tris 10 mM, pH de 8,0), seguido por centrifugación a 8.000 g. Las fracciones del sobrenadante se desechan y los gránulos de cuerpos de inclusión se almacenan de -70°C a -80°C, hasta que sean necesarios para una purificación adicional.

30 Purificación de Poliproteína

Las preparaciones de cuerpos de inclusión congeladas se descongelan a 37°C durante 15 minutos y luego se vuelven a suspender en urea 8M, NaCl 50 mM, Bis-tris-propano 20 mM, pH de 7.0 (Regulador A), empleando agitación mecánica suave. Entonces los cuerpos de inclusión re-suspendidos se agitan a temperatura ambiente con una barra de agitación magnética a 300 revoluciones por minuto durante 2 horas. Luego se centrifuga el extracto de cuerpos de inclusión a una alta velocidad y la fracción sobrenadante resultante se filtra a través de un filtro de 0,45 µM (Pall., Supor) antes del fraccionamiento cromatográfico.

El extracto de cuerpos de inclusión se aplica a una columna que contiene resina de intercambio de aniones Q Sepharose Fast Flow (QFF) (10 x 12.5 cm Amersham/Pharmacia BPG; lecho empacado en IL) previamente higienizada con NaOH 1N y luego se equilibra con el Regulador A. La columna se desarrolla a un caudal lineal de 60 centímetros/hora con el Regulador A y se recolecta el flujo atravesado que contiene predominantemente contaminantes de masa más baja para referencia. El volumen de Mtb72f se eluye en un solo paso utilizando urea 8M, NaCl 9 mM, Bis-tris-propano 20 mM, pH de 7,0 y se recolecta como un solo pico a granel basándose en la absorbancia.

Las resinas QFF son resinas de agarosa altamente reticuladas con un grupo funcional de amina cuaternaria que está positivamente cargado en las condiciones empleadas durante la purificación. La matriz cargada permite el enlace de diferentes aniones que luego se pueden eluir selectivamente utilizando un gradiente de sal. Esta cromatografía de intercambio de aniones se utiliza para separar los ácidos nucleicos y la endotoxina, que se enlazan estrechamente a la resina, de la proteína, que se enlaza más débilmente y se eluye antes que estos contaminantes. Adicionalmente, este paso remueve los contaminantes no cargados y una gran parte de las impurezas de proteína.

El pico del eluato de NaCl 90 mM es a partir de la columna de QFF y se aplica a una columna (2,6 x 12 cm Amersham/Pharmacia XK26/20; lecho empacado de 63 mililitros) conteniendo hidroxiapatita de cerámica MacroPrep® 8CHT) (tipo I, 40 uM, BioRad), previamente higienizada utilizando NaOH 1N y luego se equilibra con el Regulador C (urea 8M, NaCl 250 mM y Bis-tris-propano 20 mM, pH de 7.0). El material de flujo atravesado (FT1) que contiene la mayor parte de Mtb72f, libre de contaminantes, se recolecta. La columna se lava con el Regulador C y se recolecta cualquier material absorbente de ultravioleta resultante. Finalmente, la columna se eluye en el Regulador D (urea 8M, fosfato de sodio 200 mM, pH de 7,4).

MacroPrep® CHT es una forma macroporosa esférica de hidroxiapatita $[Ca_5(PO_4)_3OH]_2$. La cromatografía de CHT puede ser un procedimiento altamente selectivo de purificación si se encuentran condiciones de enlace y elución apropiadas. Los modos de enlace incluyen el enlace de tipo de intercambio de iones a los iones de calcio y de fosfato cargados, así como la quelación de las moléculas. El ADN se enlazarán a esta resina y se puede lograr una alta selectividad para las proteínas individuales. Las condiciones empleadas para la purificación Mtb72f sirven como un paso de pulido que permite hacer una remoción virtualmente completa de los contaminantes de la célula huésped detectables.

Durante las separaciones cromatográficas, se monitorean y se registran la absorbancia ultravioleta (UV), la conductividad, la presión, el pH, el caudal y la temperatura ambiente. El material de flujo atravesado CHT inicial (FT1) se utiliza para un procesamiento adicional corriente abajo.

Diafiltración y Filtración Estéril.

Se realiza la diafiltración sobre la reserva de CHT FT1 para remover la urea y reemplazar el regulador con Tris 20 mM, pH de 7,5. La diafiltración se lleva a cabo utilizando un sistema Pall Minim^{MR} con un dispositivo de filtración de flujo tangencial LV-Centramate^{MR}, con una membrana de ultrafiltración con un corte de peso molecular de 30 kDa (MWCO). La solución de Mtb72f en Tris 20 mM, pH de 7,5, se esteriliza en filtro utilizando un filtro esterilizante de 0,2 uM (Millipak 40). Se distribuyen 50 mililitros de la solución en botellas del medio de PETG (copolímero de tereftalato de polietileno) de 60 mililitros y luego se congelan y se almacenan a -70°C. Este material es la sustancia de fármaco a granel purificada de Mtb72f.

Ejemplo 2: Preparación de Mtb72f (seis marcas His) (SEC ID N.º:2)

Se puede seguir el procedimiento del Ejemplo 1, excepto que se omite el paso de subclonación en pPDM con el objeto de remover la marca His.

Ejemplo 3: Preparación de M72 (dos marcas His) (SEC ID N.º:4)

Construcción del Vector de Expresión M72.

El material de partida para la construcción del antígeno M72 fue el plásmido recombinante 6His-Mtb72fmut. 6His-Mtb72fmut se preparó mediante mutagénesis dirigida al sitio, utilizando el plásmido recombinante 6his-Mtb72f (véase el Ejemplo 1) como plantilla. La mutagénesis dirigida al sitio involucró el reemplazo del codón por Ser en la posición 710 en la SEC ID N.º:1, con un codón por Ala. La supresión de las cuatro histidinas N-terminales presentes en la construcción 6His-Mtb72fmut (plásmido Corixa) se logró con el "Gene Tailor Site-Directed Mutagenesis System" (Sistema de Mutagénesis Dirigida al Sitio Hecha a la Medida del Gen) (Invitrogen), conduciendo a la construcción 2His-Mtb72fmut esperada. Después de la verificación de la secuencia, se cortó la secuencia de codificación de 2His-Mtb72fmut del plásmido (mediante restricción enzimática), se purificó en gel y se ligó en el vector de expresión pET29a, dando como resultado el plásmido recombinante final pET29a/2His-Mtb72fmut. Después de la verificación de la secuencia, al plásmido recombinante se le dio la designación oficial de pRIT15497 y se utilizó para transformar las células huésped HMS174(DE3). pRIT15497 codifica para una proteína de 725 aminoácidos denominada como M72.

Producción de la Proteína M72.

Se puede emplear el mismo proceso de producción que se describe para Mtb72f (véase el Ejemplo 1), excepto que, para la producción de M72, está ausente el cloranfenicol en el medio de fermentación.

Ejemplo Biológico 1: Un modelo de ratón de un estado inactivo/latente de infección por *M. tuberculosis*

5 Con el fin de establecer un modelo de ratón de infección latente por *M. tuberculosis*, se utilizó la raza SWR. Los ratones SWR no están inmunocomprometidos, pero son deficientes para la secreción del componente de complemento C5 (véase Ooi y Colten, *Nature* (1979) 282:207-8). Los ratones SWR son incapaces de establecer un estado crónico de infección por Mtb, pero desarrollan una neumonía granulomatosa difusa caracterizada por grandes macrófagos epitelioides y espumosos con inclusiones cristaloides (gránulos derivados de neutrófilos o de eosinófilos que han sido fagocitositados), necrosis multifocal, acumulación de neutrófilos y linfocitos escasos (véase, Turner y colaboradores, *J. Submicrosc. Cytol. Pathol.* (2001) 33(1-2):217-9; y Turner y colaboradores, *Infect Immun.* (2003) 71(9):5266-72). A continuación está el protocolo para utilizar la raza de ratón Swiss Webster (SWR/J) en un modelo de infección latente por *M. tuberculosis*, con el fin de evaluar la eficacia terapéutica de Mtb72f (SEC ID N.º:6) formulada con el adyuvante AS01B. Se prepara AS01B en doble concentración mediante la adición de QS21 (5 microgramos) a vesículas unilamelares pequeñas (SUV) de dioleoil-fosfatidil-colina (100 microgramos) conteniendo colesterol (25 microgramos) (documento WO 96/33739) y monofosforil-lípido A (MPL) (5microgramos) en la membrana (véase la Publicación de Patente de los EE.UU. n.º: 2003/0143240). Se prepara una alícuota para inyección (50 microlitros) mezclando 4 microgramos de proteína en el regulador (suero regulado con fosfato, pH de 6,8), con 50 microlitros de AS01B a doble concentración. Cada ratón recibió dos inyecciones de 50 microlitros (es decir, 8 microgramos de proteína).

En la Figura 1 se ilustra esquemáticamente una línea de tiempo representativa para establecer un modelo de infección latente por *M. tuberculosis*.

Día 1: Infección mediante aerosol con 50 a 100 unidades formadoras de colonias (CFU) de organismos de *M. tuberculosis*.

25 **Días 30-90:** Tratamiento de un subconjunto de ratones con 50 miligramos de rifampina/85 miligramos de isoniazida por litro del agua para beber.

Día 61: Todos los ratones que recibieron la vacuna candidata 5, se deben inmunizar con rMtb72f+AS01B.

Día 82: Todos los ratones que recibieron la vacuna candidata se deben inmunizar con rMtb72f+AS01B.

Día 103: Todos los ratones que recibieron la vacuna candidata se deben inmunizar con rMtb72f+AS01B.

30 **Día 113:** Sangrado para los ensayos de IgG.

Diferentes Puntos del Tiempo: Se toman los bazos y los pulmones para la enumeración de unidades formadoras de colonias e inmunogenicidad.

35 **Variación 1** → Tratamiento con quimioterapia durante 60 días. Se inicia en el día 30 → se descansa durante 3, 4, 5 meses → unidades formadoras de colonias en dos ratones en cada punto del tiempo y se dejan de 4 a 7 ratones para los estudios de supervivencia.

Variación 2 → Tratamiento con quimioterapia durante 90 días. Se inicia en el día 30 → se descansa durante 4, 5 meses → unidades formadoras de colonias en dos ratones en cada punto del tiempo y se dejan 7 ratones para los estudios de supervivencia.

40 **Variación 3** → Se descansa durante 4, 5, 6 meses → unidades formadoras de colonias en dos ratones en cada punto del tiempo y se dejan 4 ratones para los estudios de supervivencia.

Variación 4 → Tratamiento con quimioterapia durante 60 días. Se inicia en el día 30 → tres inmunizaciones con r72F+AS01B intramuscularmente (i.m.) empezando en el Día 60 → se descansa durante 3, 4, 5 meses → unidades formadoras de colonias en dos ratones en cada punto del tiempo y se dejan de 4 a 7 ratones para los estudios de supervivencia.

45 **Variación 5** → Tratamiento con quimioterapia durante 90 días. Se inicia en el día 30 → tres inmunizaciones con r72F+AS01B intramuscularmente empezando en el Día 60 → se descansa durante 4, 5 meses → unidades formadoras de colonias en dos ratones en cada punto del tiempo y se dejan de 4 a 7 ratones para los estudios de supervivencia.

50 El análisis de las respuestas de anticuerpos después de la infección utilizando rMtb72f para recubrir las placas ELISA, reveló que los grupos que recibieron una combinación de quimioterapia e inmunización con Mtb72f+AS01B

tuvieron una respuesta de anticuerpos más alta (OD de hasta 2,0) que los ratones que no fueron tratados o que recibieron solamente el tratamiento de quimioterapia (OD de menos de 0,5) (Figura 2). Los ratones inmunizados con Mtb72f presentaron una respuesta de anticuerpos específica de Mtb72f dimensionable (OD de entre 1,5 y 2,5), ya fuera que recibieran quimioterapia de 60 días o bien de 90 días (Figura 3).

- 5 Las células de bazo se cosecharon de los ratones a diferentes intervalos después de que se infectaron los ratones con *M. tuberculosis*. Los esplenocitos se volvieron a estimular *in vitro* con los antígenos recombinantes para medir la secreción de IFN- γ . Los niveles de IFN- γ producidos por estas células fueron uniformemente negligibles en los grupos 1 (no tratados) y 2 (solamente quimioterapia) en el Día 60, con la excepción de Mtb39. Los estímulos de control positivo con conA, PPD y el lisado BCG, demostraron que las células eran capaces de sintetizar y secretar
- 10 IFN- γ en respuesta a otras moléculas estimulantes (Figura 4). Los niveles de IFN- γ fueron altos en los grupos que recibieron Mtb72f+AS01B, pero fueron bajos o negligibles en los grupos que no habían sido inmunizados con Mtb72f+AS01B, ya sea que hubieran o no recibido quimioterapia (Figura 5).

- 15 Durante el curso de la infección por tuberculosis y el tratamiento subsiguiente, responden las células-T específicas. Utilizando un tefido de citocina intracelular para IFN γ , se midió el porcentaje de respuestas de células CD4⁺ a Mtb72f (Figuras 6A y 6B). Pareció no haber cambios en las respuestas de células-T CD4⁺ IFN- γ ⁺ específicas para Mtb72f durante el curso de la quimioterapia sola en cualquier punto del tiempo, como se midió mediante este ensayo (Figura 7). En el día 120 después de la infección por Mtb, la tendencia de la respuesta de CD4⁺ IFN γ ⁺ a Mtb72f en los grupos que recibieron la vacuna de Mtb72f más AS01B, pareció aumentar con la duración de tiempo con quimioterapia (Figura 7).

- 20 Los resultados de nuestros experimentos demuestran que los ratones SWR son susceptibles a la infección con *M. tuberculosis*. Si se dejan sin tratamiento, los ratones SWR mueren a los 115 días después de la infección por Mtb (Figuras 8 y 9). El tiempo de supervivencia medio para los ratones que recibieron 60 días de quimioterapia de combinación fue de 170 días (Figuras 8 y 9). El tiempo de supervivencia medio para los ratones que recibieron 60 días de quimioterapia de combinación y tres inmunizaciones de Mtb72f/AS01B fue de 215 días (Figuras 8 y 9). La supervivencia para el grupo de ratones que recibieron quimioterapia es significativamente diferente (intervalo de confianza del 95 % ($p=0,0067$), de aquellos que recibieron quimioterapia y la vacuna con Mtb72f/AS01B.
- 25

LISTADO DE SECUENCIAS

<110> Coler, Rhea N Reed, Steven G Lobet, Yves

30

<120> Procedimiento novedoso para prevenir o tratar infección por *M. tuberculosis*

<130> Documento VB61507EPD1

35 <160> 6

<170> Versión de patente 3.3

<210> 1

40 <211> 2287

<212> ADN

<213> Artificial

<220>

45 <223> Descripción de secuencia artificial: proteína tri-fusión Mtb72f (Ra12-TbH9-Ra35 o Mtb32-Mtb39).

<220>

<221> rasgo_misc

<222> (30)..(30)

5 <223> n e s a , c , g o t

<220>

<221> rasgo_misc

<222> (33)..(33)

10 <223> n e s a , c , g o t

<220>

<221> CDS

<222> (42)..(2228)

15

<220>

<221> rasgo_misc

<222> (2270)..(2270)

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

20 <400> 1

tctagaaata attttgttta ctttaagaan ganatataca t atg cat cac cat cac	56
Met His His His His	
1 5	
cat cac acg gcc gcg tcc gat aac ttc cag ctg tcc cag ggt ggg cag	104
His His Thr Ala Ala Ser Asp Asn Phe Gln Leu Ser Gln Gly Gly Gln	
10 15 20	
gga ttc gcc att ccg atc ggg cag gcg atg gcg atc gcg ggc cag atc	152
Gly Phe Ala Ile Pro Ile Gly Gln Ala Met Ala Ile Ala Gly Gln Ile	
25 30 35	
cga tcg ggt ggg ggg tca ccc acc gtt cat atc ggg cct acc gcc ttc	200
Arg Ser Gly Gly Gly Ser Pro Thr Val His Ile Gly Pro Thr Ala Phe	
40 45 50	
ctc ggc ttg ggt gtt gtc gac aac aac ggc aac ggc gca cga gtc caa	248
Leu Gly Leu Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val Gln	
55 60 65	

ES 2 524 570 T3

cgc gtg gtc ggg agc gct ccg gcg gca agt ctc ggc atc tcc acc ggc Arg Val Val Gly Ser Ala Pro Ala Ala Ser Leu Gly Ile Ser Thr Gly 70 75 80 85	296
gac gtg atc acc gcg gtc gac ggc gct ccg atc aac tcg gcc acc gcg Asp Val Ile Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr Ala 90 95 100	344
atg gcg gac gcg ctt aac ggg cat cat ccc ggt gac gtc atc tcg gtg Met Ala Asp Ala Leu Asn Gly His His Pro Gly Asp Val Ile Ser Val 105 110 115	392
acc tgg caa acc aag tcg ggc ggc acg cgt aca ggg aac gtg aca ttg Thr Trp Gln Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr Leu 120 125 130	440
gcc gag gga ccc ccg gcc gaa ttc atg gtg gat ttc ggg gcg tta cca Ala Glu Gly Pro Pro Ala Glu Phe Met Val Asp Phe Gly Ala Leu Pro 135 140 145	488
ccg gag atc aac tcc gcg agg atg tac gcc ggc ccg ggt tcg gcc tcg Pro Glu Ile Asn Ser Ala Arg Met Tyr Ala Gly Pro Gly Ser Ala Ser 150 155 160 165	536
ctg gtg gcc gcg gct cag atg tgg gac agc gtg gcg agt gac ctg ttt Leu Val Ala Ala Ala Gln Met Trp Asp Ser Val Ala Ser Asp Leu Phe 170 175 180	584
tcg gcc gcg tcg gcg ttt cag tcg gtg gtc tgg ggt ctg acg gtg ggg Ser Ala Ala Ser Ala Phe Gln Ser Val Val Trp Gly Leu Thr Val Gly 185 190 195	632
tcg tgg ata ggt tcg tcg gcg ggt ctg atg gtg gcg gcg gcc tcg ccg Ser Trp Ile Gly Ser Ser Ala Gly Leu Met Val Ala Ala Ala Ser Pro 200 205 210	680
tat gtg gcg tgg atg agc gtc acc gcg ggg cag gcc gag ctg acc gcc Tyr Val Ala Trp Met Ser Val Thr Ala Gly Gln Ala Glu Leu Thr Ala 215 220 225	728
gcc cag gtc cgg gtt gct gcg gcg gcc tac gag acg gcg tat ggg ctg Ala Gln Val Arg Val Ala Ala Ala Tyr Glu Thr Ala Tyr Gly Leu 230 235 240 245	776
acg gtg ccc ccg ccg gtg atc gcc gag aac cgt gct gaa ctg atg att Thr Val Pro Pro Pro Val Ile Ala Glu Asn Arg Ala Glu Leu Met Ile 250 255 260	824
ctg ata gcg acc aac ctc ttg ggg caa aac acc ccg gcg atc gcg gtc Leu Ile Ala Thr Asn Leu Leu Gly Gln Asn Thr Pro Ala Ile Ala Val 265 270 275	872
aac gag gcc gaa tac ggc gag atg tgg gcc caa gac gcc gcc gcg atg Asn Glu Ala Glu Tyr Gly Glu Met Trp Ala Gln Asp Ala Ala Ala Met 280 285 290	920
ttt ggc tac gcc gcg gcg acg gcg acg gcg acg gcg acg ttg ctg ccg Phe Gly Tyr Ala Ala Ala Thr Ala Thr Ala Thr Ala Thr Leu Leu Pro 295 300 305	968
ttc gag gag gcg ccg gag atg acc agc gcg ggt ggg ctc ctc gag cag Phe Glu Glu Ala Pro Glu Met Thr Ser Ala Gly Gly Leu Leu Glu Gln 310 315 320 325	1016
gcc gcc gcg gtc gag gag gcc tcc gac acc gcc gcg gcg aac cag ttg	1064

ES 2 524 570 T3

Ala	Ala	Ala	Val	Glu	Glu	Ala	Ser	Asp	Thr	Ala	Ala	Ala	Asn	Gln	Leu	
				330					335					340		
atg	aac	aat	gtg	ccc	cag	gcg	ctg	caa	cag	ctg	gcc	cag	ccc	acg	cag	1112
Met	Asn	Asn	Val	Pro	Gln	Ala	Leu	Gln	Gln	Leu	Ala	Gln	Pro	Thr	Gln	
			345					350					355			
ggc	acc	acg	cct	tct	tcc	aag	ctg	ggt	ggc	ctg	tgg	aag	acg	gtc	tcg	1160
Gly	Thr	Thr	Pro	Ser	Ser	Lys	Leu	Gly	Gly	Leu	Trp	Lys	Thr	Val	Ser	
			360				365						370			
ccg	cat	cgg	tcg	ccg	atc	agc	aac	atg	gtg	tcg	atg	gcc	aac	aac	cac	1208
Pro	His	Arg	Ser	Pro	Ile	Ser	Asn	Met	Val	Ser	Met	Ala	Asn	Asn	His	
	375					380					385					
atg	tcg	atg	acc	aac	tcg	ggt	gtg	tcg	atg	acc	aac	acc	ttg	agc	tcg	1256
Met	Ser	Met	Thr	Asn	Ser	Gly	Val	Ser	Met	Thr	Asn	Thr	Leu	Ser	Ser	
					395					400					405	
atg	ttg	aag	ggc	ttt	gct	ccg	gcg	gcg	gcc	gcc	cag	gcc	gtg	caa	acc	1304
Met	Leu	Lys	Gly	Phe	Ala	Pro	Ala	Ala	Ala	Ala	Gln	Ala	Val	Gln	Thr	
				410					415					420		
gcg	gcg	caa	aac	ggg	gtc	cgg	gcg	atg	agc	tcg	ctg	ggc	agc	tcg	ctg	1352
Ala	Ala	Gln	Asn	Gly	Val	Arg	Ala	Met	Ser	Ser	Leu	Gly	Ser	Ser	Leu	
			425					430					435			
ggt	tct	tcg	ggt	ctg	ggc	ggt	ggg	gtg	gcc	gcc	aac	ttg	ggt	cgg	gcg	1400
Gly	Ser	Ser	Gly	Leu	Gly	Gly	Gly	Val	Ala	Ala	Asn	Leu	Gly	Arg	Ala	
			440				445					450				
gcc	tcg	gtc	ggt	tcg	ttg	tcg	gtg	ccg	cag	gcc	tgg	gcc	gcg	gcc	aac	1448
Ala	Ser	Val	Gly	Ser	Leu	Ser	Val	Pro	Gln	Ala	Trp	Ala	Ala	Ala	Asn	
			455			460					465					
cag	gca	gtc	acc	ccg	gcg	gcg	cgg	gcg	ctg	ccg	ctg	acc	agc	ctg	acc	1496
Gln	Ala	Val	Thr	Pro	Ala	Ala	Arg	Ala	Leu	Pro	Leu	Thr	Ser	Leu	Thr	
					475					480					485	
agc	gcc	gcg	gaa	aga	ggg	ccc	ggg	cag	atg	ctg	ggc	ggg	ctg	ccg	gtg	1544
Ser	Ala	Ala	Glu	Arg	Gly	Pro	Gly	Gln	Met	Leu	Gly	Gly	Leu	Pro	Val	
				490					495					500		
ggg	cag	atg	ggc	gcc	agg	gcc	ggt	ggt	ggg	ctc	agt	ggt	gtg	ctg	cgt	1592
Gly	Gln	Met	Gly	Ala	Arg	Ala	Gly	Gly	Gly	Leu	Ser	Gly	Val	Leu	Arg	
				505				510					515			
ggt	ccg	ccg	cga	ccc	tat	gtg	atg	ccg	cat	tct	ccg	gca	gcc	ggc	gat	1640
Val	Pro	Pro	Arg	Pro	Tyr	Val	Met	Pro	His	Ser	Pro	Ala	Ala	Gly	Asp	
				520			525					530				
atc	gcc	ccg	ccg	gcc	ttg	tcg	cag	gac	cgg	ttc	gcc	gac	ttc	ccc	gcg	1688
Ile	Ala	Pro	Pro	Ala	Leu	Ser	Gln	Asp	Arg	Phe	Ala	Asp	Phe	Pro	Ala	
				535			540				545					
ctg	ccc	ctc	gac	ccg	tcc	gcg	atg	gtc	gcc	caa	gtg	ggg	cca	cag	gtg	1736
Leu	Pro	Leu	Asp	Pro	Ser	Ala	Met	Val	Ala	Gln	Val	Gly	Pro	Gln	Val	
					555					560					565	
gtc	aac	atc	aac	acc	aaa	ctg	ggc	tac	aac	aac	gcc	gtg	ggc	gcc	ggg	1784
Val	Asn	Ile	Asn	Thr	Lys	Leu	Gly	Tyr	Asn	Asn	Ala	Val	Gly	Ala	Gly	
				570					575					580		
acc	ggc	atc	gtc	atc	gat	ccc	aac	ggt	gtc	gtg	ctg	acc	aac	aac	cac	1832
Thr	Gly	Ile	Val	Ile	Asp	Pro	Asn	Gly	Val	Val	Leu	Thr	Asn	Asn	His	

ES 2 524 570 T3

	585	590	595	
	gtg atc gcg ggc gcc acc gac atc aat gcg ttc agc gtc ggc tcc ggc			1880
	Val Ile Ala Gly Ala Thr Asp Ile Asn Ala Phe Ser Val Gly Ser Gly			
	600	605	610	
	caa acc tac ggc gtc gat gtg gtc ggg tat gac cgc acc cag gat gtc			1928
	Gln Thr Tyr Gly Val Asp Val Val Gly Tyr Asp Arg Thr Gln Asp Val			
	615	620	625	
	gcg gtg ctg cag ctg cgc ggt gcc ggt ggc ctg ccg tcg gcg gcg atc			1976
	Ala Val Leu Gln Leu Arg Gly Ala Gly Gly Leu Pro Ser Ala Ala Ile			
	630	635	640	645
	ggt ggc ggc gtc gcg gtt ggt gag ccc gtc gtc gcg atg ggc aac agc			2024
	Gly Gly Gly Val Ala Val Gly Glu Pro Val Val Ala Met Gly Asn Ser			
	650	655	660	
	ggt ggg cag ggc gga acg ccc cgt gcg gtg cct ggc agg gtg gtc gcg			2072
	Gly Gly Gln Gly Gly Thr Pro Arg Ala Val Pro Gly Arg Val Val Ala			
	665	670	675	
	ctc ggc caa acc gtg cag gcg tcg gat tcg ctg acc ggt gcc gaa gag			2120
	Leu Gly Gln Thr Val Gln Ala Ser Asp Ser Leu Thr Gly Ala Glu Glu			
	680	685	690	
	aca ttg aac ggg ttg atc cag ttc gat gcc gcg atc cag ccc ggt gat			2168
	Thr Leu Asn Gly Leu Ile Gln Phe Asp Ala Ala Ile Gln Pro Gly Asp			
	695	700	705	
	tcg ggc ggg ccc gtc gtc aac ggc cta gga cag gtg gtc ggt atg aac			2216
	Ser Gly Gly Pro Val Val Asn Gly Leu Gly Gln Val Val Gly Met Asn			
	710	715	720	725
	acg gcc gcg tcc taggatatcc atcacactgg cggccgctcg agcagatccg			2268
	Thr Ala Ala Ser			
	gntgtaacaa agcccgaaa			2287

<210> 2

<211> 729

<212> PRT

5 <213> Artificial

<220>

<223> Descripción de secuencia artificial: proteína tri-fusión Mtb72f (Ra12-TbH9-Ra35 o Mtb32-Mtb39).

10 <400> 2

ES 2 524 570 T3

Met His His His His His His Thr Ala Ala Ser Asp Asn Phe Gln Leu
 1 5 10 15

Ser Gln Gly Gly Gln Gly Phe Ala Ile Pro Ile Gly Gln Ala Met Ala
 20 25 30

Ile Ala Gly Gln Ile Arg Ser Gly Gly Gly Ser Pro Thr Val His Ile
 35 40 45

Gly Pro Thr Ala Phe Leu Gly Leu Gly Val Val Asp Asn Asn Gly Asn
 50 55 60

Gly Ala Arg Val Gln Arg Val Val Gly Ser Ala Pro Ala Ala Ser Leu
 65 70 75 80

Gly Ile Ser Thr Gly Asp Val Ile Thr Ala Val Asp Gly Ala Pro Ile
 85 90 95

Asn Ser Ala Thr Ala Met Ala Asp Ala Leu Asn Gly His His Pro Gly
 100 105 110

Asp Val Ile Ser Val Thr Trp Gln Thr Lys Ser Gly Gly Thr Arg Thr
 115 120 125

Gly Asn Val Thr Leu Ala Glu Gly Pro Pro Ala Glu Phe Met Val Asp
 130 135 140

ES 2 524 570 T3

Phe Gly Ala Leu Pro Pro Glu Ile Asn Ser Ala Arg Met Tyr Ala Gly
 145 150 155 160
 Pro Gly Ser Ala Ser Leu Val Ala Ala Ala Gln Met Trp Asp Ser Val
 165 170 175
 Ala Ser Asp Leu Phe Ser Ala Ala Ser Ala Phe Gln Ser Val Val Trp
 180 185 190
 Gly Leu Thr Val Gly Ser Trp Ile Gly Ser Ser Ala Gly Leu Met Val
 195 200 205
 Ala Ala Ala Ser Pro Tyr Val Ala Trp Met Ser Val Thr Ala Gly Gln
 210 215 220
 Ala Glu Leu Thr Ala Ala Gln Val Arg Val Ala Ala Ala Ala Tyr Glu
 225 230 235 240
 Thr Ala Tyr Gly Leu Thr Val Pro Pro Pro Val Ile Ala Glu Asn Arg
 245 250 255
 Ala Glu Leu Met Ile Leu Ile Ala Thr Asn Leu Leu Gly Gln Asn Thr
 260 265 270
 Pro Ala Ile Ala Val Asn Glu Ala Glu Tyr Gly Glu Met Trp Ala Gln
 275 280 285
 Asp Ala Ala Ala Met Phe Gly Tyr Ala Ala Ala Thr Ala Thr Ala Thr
 290 295 300
 Ala Thr Leu Leu Pro Phe Glu Glu Ala Pro Glu Met Thr Ser Ala Gly
 305 310 315 320
 Gly Leu Leu Glu Gln Ala Ala Ala Val Glu Glu Ala Ser Asp Thr Ala
 325 330 335
 Ala Ala Asn Gln Leu Met Asn Asn Val Pro Gln Ala Leu Gln Gln Leu
 340 345 350
 Ala Gln Pro Thr Gln Gly Thr Thr Pro Ser Ser Lys Leu Gly Gly Leu

ES 2 524 570 T3

	355		360		365														
Trp	Lys	Thr	Val	Ser	Pro	His	Arg	Ser	Pro	Ile	Ser	Asn	Met	Val	Ser				
	370					375					380								
Met	Ala	Asn	Asn	His	Met	Ser	Met	Thr	Asn	Ser	Gly	Val	Ser	Met	Thr				
	385				390					395					400				
Asn	Thr	Leu	Ser	Ser	Met	Leu	Lys	Gly	Phe	Ala	Pro	Ala	Ala	Ala	Ala				
				405					410					415					
Gln	Ala	Val	Gln	Thr	Ala	Ala	Gln	Asn	Gly	Val	Arg	Ala	Met	Ser	Ser				
			420					425					430						
Leu	Gly	Ser	Ser	Leu	Gly	Ser	Ser	Gly	Leu	Gly	Gly	Gly	Val	Ala	Ala				
		435					440						445						
Asn	Leu	Gly	Arg	Ala	Ala	Ser	Val	Gly	Ser	Leu	Ser	Val	Pro	Gln	Ala				
	450					455						460							
Trp	Ala	Ala	Ala	Asn	Gln	Ala	Val	Thr	Pro	Ala	Ala	Arg	Ala	Leu	Pro				
	465				470					475					480				
Leu	Thr	Ser	Leu	Thr	Ser	Ala	Ala	Glu	Arg	Gly	Pro	Gly	Gln	Met	Leu				
				485					490					495					
Gly	Gly	Leu	Pro	Val	Gly	Gln	Met	Gly	Ala	Arg	Ala	Gly	Gly	Gly	Leu				
			500					505					510						
Ser	Gly	Val	Leu	Arg	Val	Pro	Pro	Arg	Pro	Tyr	Val	Met	Pro	His	Ser				
		515					520					525							
Pro	Ala	Ala	Gly	Asp	Ile	Ala	Pro	Pro	Ala	Leu	Ser	Gln	Asp	Arg	Phe				
	530					535						540							
Ala	Asp	Phe	Pro	Ala	Leu	Pro	Leu	Asp	Pro	Ser	Ala	Met	Val	Ala	Gln				
	545				550					555					560				
Val	Gly	Pro	Gln	Val	Val	Asn	Ile	Asn	Thr	Lys	Leu	Gly	Tyr	Asn	Asn				
				565					570					575					

ES 2 524 570 T3

Ala Val Gly Ala Gly Thr Gly Ile Val Ile Asp Pro Asn Gly Val Val
580 585 590

Leu Thr Asn Asn His Val Ile Ala Gly Ala Thr Asp Ile Asn Ala Phe
595 600 605

Ser Val Gly Ser Gly Gln Thr Tyr Gly Val Asp Val Val Gly Tyr Asp
610 615 620

Arg Thr Gln Asp Val Ala Val Leu Gln Leu Arg Gly Ala Gly Gly Leu
625 630 635 640

Pro Ser Ala Ala Ile Gly Gly Gly Val Ala Val Gly Glu Pro Val Val
645 650 655

Ala Met Gly Asn Ser Gly Gly Gln Gly Gly Thr Pro Arg Ala Val Pro
660 665 670

Gly Arg Val Val Ala Leu Gly Gln Thr Val Gln Ala Ser Asp Ser Leu
675 680 685

Thr Gly Ala Glu Glu Thr Leu Asn Gly Leu Ile Gln Phe Asp Ala Ala
690 695 700

Ile Gln Pro Gly Asp Ser Gly Gly Pro Val Val Asn Gly Leu Gly Gln
705 710 715 720

Val Val Gly Met Asn Thr Ala Ala Ser
725

<210> 3

<211> 2178

<212> ADN

5 <213> Artificial

<220>

<223> Descripción de secuencia artificial: variante de proteína tri-fusión Mtb72f (Ra12-TbH9-Ra35 o Mtb32-Mtb39).

<220>

10 <221> rasgo_misc

<222> (4)..(9)

<223> "cat cac" en lugar de "cat cac cat cac cat cac"

<220>

5 <221> rasgo_misc

<222> (2116)..(2118)

<223> "gcg" en lugar de "tcg"

<400> 3

10 **atgcatcaca cggccgcgtc cgataacttc cagctgtccc aggggtgggca gggattcgcc 60**
 attccgatcg ggcaggcgat ggcgatcgcg ggccagatcc gatcgggtgg ggggtcaccc 120

ES 2 524 570 T3

accgttcata tcgggcctac cgccttcctc ggcttgggtg ttgtcgacaa caacggcaac 180
 ggcgcacgag tccaacgcgt ggtcgggagc gctccggcgg caagtctcgg catctccacc 240
 ggcgacgtga tcaccgcggt cgacggcgct ccgatcaact cggccaccgc gatggcggac 300
 gcgcttaacg ggcatcatcc cggtgacgtc atctcgggtga cctggcaaac caagtcgggc 360
 ggcacgcgta cagggaacgt gacattggcc gagggacccc cggccgaatt catggtggat 420
 ttcggggcgt taccaccgga gatcaactcc gcgaggatgt acgccggccc gggttcggcc 480
 tcgctgggtg ccgcggctca gatgtgggac agcgtggcga gtgacctgtt ttcggccgcg 540
 tcggcgtttc agtcggtggt ctggggtctg acggtggggt cgtggatagg ttcgtcggcg 600
 ggtctgatgg tggcggcggc ctcgccgat gtggcgtgga tgagcgtcac cgcggggcag 660
 gccgagctga ccgccgccca ggtccgggtt gctgcggcgg cctacgagac ggcgatggg 720
 ctgacgggtc ccccgccggt gatcgccgag aaccgtgctg aactgatgat tctgatagcg 780
 accaacctct tggggcaaaa caccgccggc atcgcggtca acgaggccga atacggcgag 840
 atgtgggccc aagacgccgc cgcgatgttt ggctacgccg cggcgacggc gacggcgacg 900
 gcgacgttgc tgccgttcga ggaggcgccg gagatgacca gcgcgggtgg gtcctcagag 960
 caggccgccc cggtcgagga ggcctccgac accgcccgcg cgaaccagtt gatgaacaat 1020
 gtgccccagg cgctgcaaca gctggcccag cccacgcagg gcaccacgcc ttcttccaag 1080
 ctgggtggcc tgtggaagac ggtctcgccg catcggtcgc cgatcagcaa catggtgtcg 1140
 atggccaaca accacatgtc gatgaccaac tcgggtgtgt cgatgaccaa caccttgagc 1200
 tcgatgttga agggctttgc tccggcggcg gccgcccagg ccgtgcaaac cgcggcgcaa 1260
 aacgggggtcc gggcgatgag ctcgctgggc agctcgetgg gttcttcggg tctgggcggg 1320
 ggggtggccc ccaacttggg tcgggcggcc tcggtcgggt cgttgteggg gccgcaggcc 1380
 tgggccgccc ccaaccaggc agtcaccccc gcggcgcggg cgctgccgct gaccagcctg 1440
 accagcggcg cggaaagagg gcccgggcag atgctgggcg ggctgccggg ggggcagatg 1500
 ggcgccaggg ccggtggtgg gctcagtggt gtgctgcgtg ttccgccgcg accctatgtg 1560
 atgccgcatt ctccggcagc cggcgatata gcccgcggc ccttgtegca ggaccggttc 1620
 gccgacttcc ccgcgctgcc cctcgacccc tccgcgatgg tcgcccagt ggggccacag 1680
 gtggtcaaca tcaacaccaa actgggctac aacaacgccg tgggcgcccg gaccggcatc 1740
 gtcacgatc ccaacggtgt cgtgctgacc aacaaccacg tgatcgcggg cgccaccgac 1800
 atcaatgctg tcagcgtcgg ctccggccaa acctacggcg tcgatgtggt cgggtatgac 1860
 cgcacccagg atgtcgcggt gctgcagctg cgcggtgccg gtggcctgcc gtcggcggcg 1920
 atcgggtggcg gcgtcgcggt tgggtgagccc gtcgtcgcga tgggcaacag cgggtggcag 1980
 ggcggaacgc cccgtgcggt gcctggcagg gtggtcgcgc tcggccaaac cgtgcaggcg 2040

ES 2 524 570 T3

tcgattcgc **t**gaccggtgc **c**gaagagaca **t**tgaacgggt **t**gatccagtt **c**gatgccgcg **2100**
atccagcccc **g**tgatgcggg **c**gggcccgtc **g**tcaacggcc **t**aggacaggt **g**gtcggtatg **2160**
aacacggccg **c**gtcctag **2178**

<210> 4

<211> 725

<212> PRT

5 <213> Artificial

<220>

<223> Descripción de secuencia artificial: variante de proteína tri-fusión Mtb72f (Ra12-TbH9-Ra35 o Mtb32-Mtb39).

10 <220>

<221> rasgo_misc

<222> (2)..(3)

<223> "His His" en lugar de "His His His His His His"

15 <220>

<221> rasgo_misc

<222> (706)..(706)

<223> "Ala" en lugar de "Ser"

20 <400> 4

ES 2 524 570 T3

Met His His Thr Ala Ala Ser Asp Asn Phe Gln Leu Ser Gln Gly Gly
1 5 10 15

Gln Gly Phe Ala Ile Pro Ile Gly Gln Ala Met Ala Ile Ala Gly Gln
 20 25 30

Ile Arg Ser Gly Gly Gly Ser Pro Thr Val His Ile Gly Pro Thr Ala
 35 40 45

Phe Leu Gly Leu Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val
 50 55 60

Gln Arg Val Val Gly Ser Ala Pro Ala Ala Ser Leu Gly Ile Ser Thr
65 70 75 80

Gly Asp Val Ile Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr
 85 90 95

ES 2 524 570 T3

Ala Met Ala Asp Ala Leu Asn Gly His His Pro Gly Asp Val Ile Ser
100 105 110

Val Thr Trp Gln Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr
115 120 125

Leu Ala Glu Gly Pro Pro Ala Glu Phe Met Val Asp Phe Gly Ala Leu
130 135 140

Pro Pro Glu Ile Asn Ser Ala Arg Met Tyr Ala Gly Pro Gly Ser Ala
145 150 155 160

Ser Leu Val Ala Ala Ala Gln Met Trp Asp Ser Val Ala Ser Asp Leu
165 170 175

Phe Ser Ala Ala Ser Ala Phe Gln Ser Val Val Trp Gly Leu Thr Val
180 185 190

Gly Ser Trp Ile Gly Ser Ser Ala Gly Leu Met Val Ala Ala Ala Ser
195 200 205

Pro Tyr Val Ala Trp Met Ser Val Thr Ala Gly Gln Ala Glu Leu Thr
210 215 220

Ala Ala Gln Val Arg Val Ala Ala Ala Ala Tyr Glu Thr Ala Tyr Gly
225 230 235 240

Leu Thr Val Pro Pro Pro Val Ile Ala Glu Asn Arg Ala Glu Leu Met
245 250 255

Ile Leu Ile Ala Thr Asn Leu Leu Gly Gln Asn Thr Pro Ala Ile Ala
260 265 270

Val Asn Glu Ala Glu Tyr Gly Glu Met Trp Ala Gln Asp Ala Ala Ala
275 280 285

Met Phe Gly Tyr Ala Ala Ala Thr Ala Thr Ala Thr Ala Thr Leu Leu
290 295 300

Pro Phe Glu Glu Ala Pro Glu Met Thr Ser Ala Gly Gly Leu Leu Glu

ES 2 524 570 T3

305		310		315		320
Gln Ala Ala Ala Val	Glu Glu Ala Ser Asp Thr Ala Ala Ala	Asn Gln				
	325		330			335
Leu Met Asn Asn Val	Pro Gln Ala Leu Gln Gln Leu Ala Gln Pro Thr					
	340		345			350
Gln Gly Thr Thr Pro Ser Ser Lys Leu Gly Gly Leu Trp Lys Thr Val						
	355		360			365
Ser Pro His Arg Ser Pro Ile Ser Asn Met Val Ser Met Ala Asn Asn						
	370		375			380
His Met Ser Met Thr Asn Ser Gly Val Ser Met Thr Asn Thr Leu Ser						
	385		390			395
Ser Met Leu Lys Gly Phe Ala Pro Ala Ala Ala Ala Gln Ala Val Gln						
	405		410			415
Thr Ala Ala Gln Asn Gly Val Arg Ala Met Ser Ser Leu Gly Ser Ser						
	420		425			430
Leu Gly Ser Ser Gly Leu Gly Gly Gly Val Ala Ala Asn Leu Gly Arg						
	435		440			445
Ala Ala Ser Val Gly Ser Leu Ser Val Pro Gln Ala Trp Ala Ala Ala						
	450		455			460
Asn Gln Ala Val Thr Pro Ala Ala Arg Ala Leu Pro Leu Thr Ser Leu						
	465		470			475
Thr Ser Ala Ala Glu Arg Gly Pro Gly Gln Met Leu Gly Gly Leu Pro						
	485		490			495
Val Gly Gln Met Gly Ala Arg Ala Gly Gly Gly Leu Ser Gly Val Leu						
	500		505			510
Arg Val Pro Pro Arg Pro Tyr Val Met Pro His Ser Pro Ala Ala Gly						
	515		520			525

ES 2 524 570 T3

Asp Ile Ala Pro Pro Ala Leu Ser Gln Asp Arg Phe Ala Asp Phe Pro
 530 535 540
 Ala Leu Pro Leu Asp Pro Ser Ala Met Val Ala Gln Val Gly Pro Gln
 545 550 555 560
 Val Val Asn Ile Asn Thr Lys Leu Gly Tyr Asn Asn Ala Val Gly Ala
 565 570 575
 Gly Thr Gly Ile Val Ile Asp Pro Asn Gly Val Val Leu Thr Asn Asn
 580 585 590
 His Val Ile Ala Gly Ala Thr Asp Ile Asn Ala Phe Ser Val Gly Ser
 595 600 605
 Gly Gln Thr Tyr Gly Val Asp Val Val Gly Tyr Asp Arg Thr Gln Asp
 610 615 620
 Val Ala Val Leu Gln Leu Arg Gly Ala Gly Gly Leu Pro Ser Ala Ala
 625 630 635 640
 Ile Gly Gly Gly Val Ala Val Gly Glu Pro Val Val Ala Met Gly Asn
 645 650 655
 Ser Gly Gly Gln Gly Gly Thr Pro Arg Ala Val Pro Gly Arg Val Val
 660 665 670
 Ala Leu Gly Gln Thr Val Gln Ala Ser Asp Ser Leu Thr Gly Ala Glu
 675 680 685
 Glu Thr Leu Asn Gly Leu Ile Gln Phe Asp Ala Ala Ile Gln Pro Gly
 690 695 700
 Asp Ala Gly Gly Pro Val Val Asn Gly Leu Gly Gln Val Val Gly Met
 705 710 715 720
 Asn Thr Ala Ala Ser
 725

<210> 5

<211> 2172

<212> ADN

5 <213> Artificial

ES 2 524 570 T3

<220>

<221> rasgo_misc

<222> (4)..(4)

<223> supresión de la marca "cat cac cat cac cat cac"

5

<400> 5

```
atgacggccg cgtccgataa cttccagctg tcccagggtg ggcagggatt cgccattccg      60
atcgggcagg cgatggcgat cgcgggccag atccgatcgg gtgggggggtc acccaccggt      120
catatcgggc ctaccgcctt cctcggcttg ggtggtgtcg acaacaacgg caacggcgca      180
cgagtccaac gcgtggtcgg gagcgctccg gcggaagtc tcggcatctc caccggcgac      240
gtgatcaccg cggtcgacgg cgctccgatc aactcggcca ccgcatggc ggacgcgctt      300
aacgggcata atcccgggta cgtcatctcg gtgacctggc aaaccaagtc gggcggcacg      360
cgtacagggg acgtgacatt ggccgagggg cccccggccg aattcatggt ggatttcggg      420
gcgttaccac cggagatcaa ctccgcgagg atgtacgccg gcccgggttc ggctctgctg      480
gtggccgcgg ctcatatgtg ggacagcgtg gcgagtgacc tgttttcggc cgcgtcggcg      540
tttcagtcgg tggctcgggg tctgacgggt gggtcgtgga taggttcgtc ggcgggtctg      600
atggtggcgg cggcctcgcc gtatgtggcg tggatgagcg tcaccgcggg gcaggccgag      660
ctgaccgccg cccaggtccg ggttgctgcg gcgccctacg agacggcgta tgggctgacg      720
gtgccccgcg cggtgatcgc cgagaaccgt gctgaactga tgattctgat agcgaccaac      780
ctcttggggc aaaacacccc ggcgatcgcg gtcaacgagg ccgaatacgg cgagatgtgg      840
gccaagacg ccgccgcgat gtttggttac gccgcggcga cggcgacggc gacggcgacg      900
ttgctgccgt tcgaggagge gccggagatg accagcgcgg gtgggctcct cgagcaggcc      960
```

ES 2 524 570 T3

gccgcggtcg aggaggcctc cgacaccgcc gcggcgaacc agttgatgaa caatgtgccc 1020
 caggcgctgc aacagctggc ccagcccacg cagggcacca cgccttcttc caagctgggt 1080
 ggctgtgga agacggtctc gccgcacggtc tcgccgatca gcaacatggt gtcgatggcc 1140
 aacaaccaca tgtcgatgac caactcgggt gtgtcgatga ccaacacctt gagctcgatg 1200
 ttgaagggct ttgctccggc ggcgccgcc caggccgtgc aaaccgcggc gcaaaacggg 1260
 gtccggggcga tgagctcgtc gggcagctcg ctgggttctt cgggtctggg cgggtgggtg 1320
 gccgccaact tgggtcgggc ggctcggtc ggttcggtgt cggtgccgca ggctgggccc 1380
 gcggccaacc aggcagtcac cccggcgggc cgggcgctgc cgctgaccag cctgaccacg 1440
 gccgcggaag gagggcccgg gcagatgctg ggcgggctgc cgggtggggca gatgggccc 1500
 agggccggtg gtgggctcag tgggtgctg cgtgttccgc cgcgacccta tgtgatgccg 1560
 cattctccgg cagccggcga tctcggccc cggccttgt cgcaggaccg gttcggccac 1620
 ttccccgcgc tccccctcga cccgtccgcg atggctgccc aagtggggcc acaggtggtc 1680
 aacatcaaca ccaaactggg ctacaacaac gccgtgggcg ccgggaccgg catcgatc 1740
 gatcccaacg gtgtcgtgct gaccaacaac cacgtgatcg cgggcgccac cgacatcaat 1800
 gcgttcagcg tcggctccgg ccaaacctac ggcgtcgatg tggtcgggta tgaccgcacc 1860
 caggatgtcg cgggtctgca gctgcgcggt gccgggtggc tgccgtcggc ggcgatcggg 1920
 ggcggcgtcg cggttgggta gcccgtcgtc gcgatgggca acagcgggtg gcagggcgga 1980
 acgccccgtg cgggtcctgg caggggtggtc gcgctcggcc aaaccgtgca ggcgtcggat 2040
 tcgctgaccg gtgccgaaga gacattgaac gggttgatcc agttcgatgc cgcgatccag 2100
 cccggtgatt cgggcggggc cgtcgtcaac ggcctaggac aggtggtcgg tatgaacacg 2160
 gccgcgtcct ga 2172

<210> 6

<211> 723

<212> PRT

5 <213> Artificial

<220>

<223> Descripción de la secuencia artificial: Mtb72F-IND

10 <220>

<221> rasgo_misc

<222> (2)..(2)

<223> supresión de la marca "His His His His His His"

<400> 6

```

Met Thr Ala Ala Ser Asp Asn Phe Gln Leu Ser Gln Gly Gly Gln Gly
1           5           10           15

Phe Ala Ile Pro Ile Gly Gln Ala Met Ala Ile Ala Gly Gln Ile Arg
          20           25           30

Ser Gly Gly Gly Ser Pro Thr Val His Ile Gly Pro Thr Ala Phe Leu
          35           40           45

Gly Leu Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val Gln Arg
50           55           60

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

Val Ile Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr Ala Met
          85           90           95
    
```

ES 2 524 570 T3

Ala Asp Ala Leu Asn Gly His His Pro Gly Asp Val Ile Ser Val Thr
100 105 110

Trp Gln Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr Leu Ala
115 120 125

Glu Gly Pro Pro Ala Glu Phe Met Val Asp Phe Gly Ala Leu Pro Pro
130 135 140

Glu Ile Asn Ser Ala Arg Met Tyr Ala Gly Pro Gly Ser Ala Ser Leu
145 150 155 160

Val Ala Ala Ala Gln Met Trp Asp Ser Val Ala Ser Asp Leu Phe Ser
165 170 175

Ala Ala Ser Ala Phe Gln Ser Val Val Trp Gly Leu Thr Val Gly Ser
180 185 190

Trp Ile Gly Ser Ser Ala Gly Leu Met Val Ala Ala Ser Pro Tyr
195 200 205

Val Ala Trp Met Ser Val Thr Ala Gly Gln Ala Glu Leu Thr Ala Ala
210 215 220

Gln Val Arg Val Ala Ala Ala Tyr Glu Thr Ala Tyr Gly Leu Thr
225 230 235 240

Val Pro Pro Pro Val Ile Ala Glu Asn Arg Ala Glu Leu Met Ile Leu
245 250 255

Ile Ala Thr Asn Leu Leu Gly Gln Asn Thr Pro Ala Ile Ala Val Asn
260 265 270

Glu Ala Glu Tyr Gly Glu Met Trp Ala Gln Asp Ala Ala Ala Met Phe
275 280 285

Gly Tyr Ala Ala Ala Thr Ala Thr Ala Thr Ala Thr Leu Leu Pro Phe
290 295 300

Glu Glu Ala Pro Glu Met Thr Ser Ala Gly Gly Leu Leu Glu Gln Ala
305 310 315 320

ES 2 524 570 T3

Ala Ala Val Glu Glu Ala Ser Asp Thr Ala Ala Ala Asn Gln Leu Met
 325 330 335

Asn Asn Val Pro Gln Ala Leu Gln Gln Leu Ala Gln Pro Thr Gln Gly
 340 345 350

Thr Thr Pro Ser Ser Lys Leu Gly Gly Leu Trp Lys Thr Val Ser Pro
 355 360 365

His Arg Ser Pro Ile Ser Asn Met Val Ser Met Ala Asn Asn His Met
 370 375 380

Ser Met Thr Asn Ser Gly Val Ser Met Thr Asn Thr Leu Ser Ser Met
 385 390 395 400

Leu Lys Gly Phe Ala Pro Ala Ala Ala Ala Gln Ala Val Gln Thr Ala
 405 410 415

Ala Gln Asn Gly Val Arg Ala Met Ser Ser Leu Gly Ser Ser Leu Gly
 420 425 430

Ser Ser Gly Leu Gly Gly Gly Val Ala Ala Asn Leu Gly Arg Ala Ala
 435 440 445

Ser Val Gly Ser Leu Ser Val Pro Gln Ala Trp Ala Ala Ala Asn Gln
 450 455 460

Ala Val Thr Pro Ala Ala Arg Ala Leu Pro Leu Thr Ser Leu Thr Ser
 465 470 475 480

Ala Ala Glu Arg Gly Pro Gly Gln Met Leu Gly Gly Leu Pro Val Gly
 485 490 495

Gln Met Gly Ala Arg Ala Gly Gly Gly Leu Ser Gly Val Leu Arg Val
 500 505 510

Pro Pro Arg Pro Tyr Val Met Pro His Ser Pro Ala Ala Gly Asp Ile
 515 520 525

Ala Pro Pro Ala Leu Ser Gln Asp Arg Phe Ala Asp Phe Pro Ala Leu

REIVINDICACIONES

- 1.- Un polipéptido que comprende la secuencia de aminoácidos de SEC ID N.º: 4.
- 2.- El polipéptido de acuerdo con la reivindicación 1, que consiste en la secuencia de aminoácidos de SEC ID N.º: 4.
- 5 3.- Un polinucleótido que comprende una secuencia de ácidos nucleicos que codifica la secuencia de aminoácidos de SEC ID N.º: 4.
- 4.- El polinucleótido de acuerdo con la reivindicación 3, que comprende la secuencia de ácidos nucleicos de SEC ID N.º: 3.
- 5.- Una composición farmacéutica que comprende un polipéptido de acuerdo con la reivindicación 1.
- 10 6.- La composición farmacéutica de acuerdo con la reivindicación 5, que comprende adicionalmente 3D-MPL y QS21 en una formulación de liposomas o 3D-MPL y QS21 en una emulsión de aceite en agua.
- 7.- Una composición farmacéutica que comprende un polinucleótido de acuerdo con la reivindicación 3.
- 8.- La composición farmacéutica de acuerdo con la reivindicación 7, en la que el polinucleótido se proporciona en un vector viral.
- 15 9.- La composición farmacéutica de acuerdo con la reivindicación 7, en la que el polinucleótido se proporciona en una célula huésped bacteriana.
- 10.- La composición farmacéutica de acuerdo con la reivindicación 9, en la que la bacteria es *Bacillus Calmette-Guerin*.

Figura 1

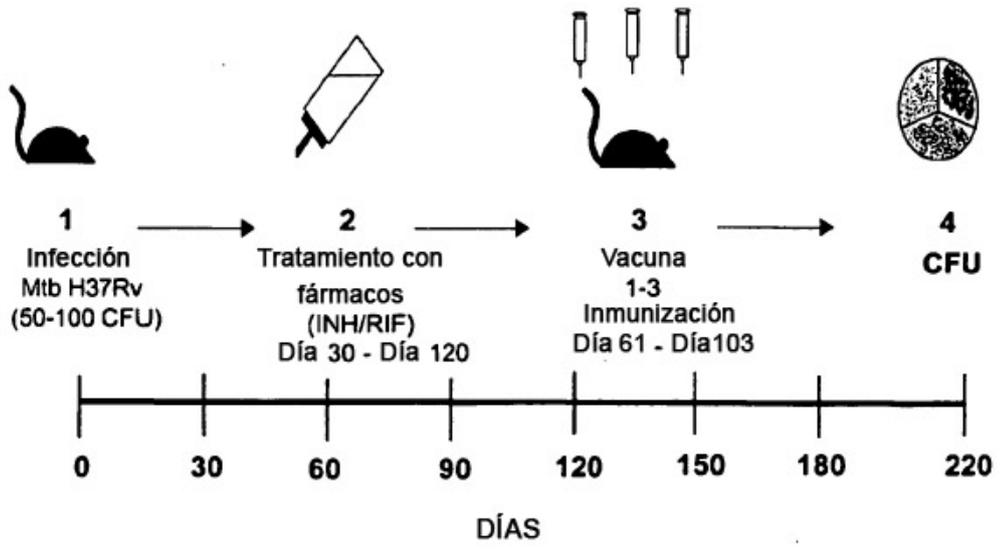


Figura 2

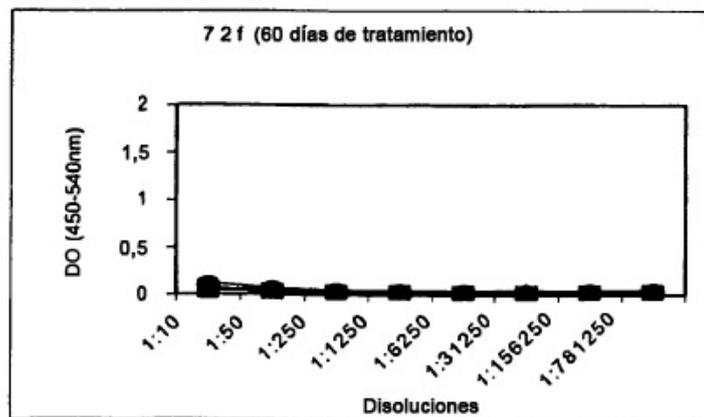
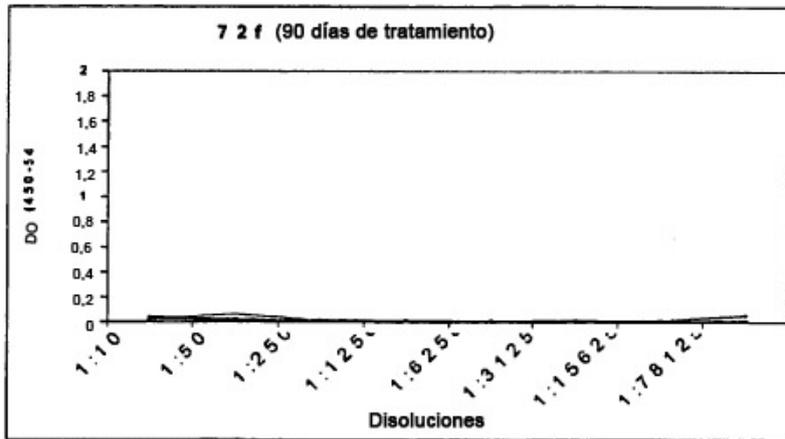
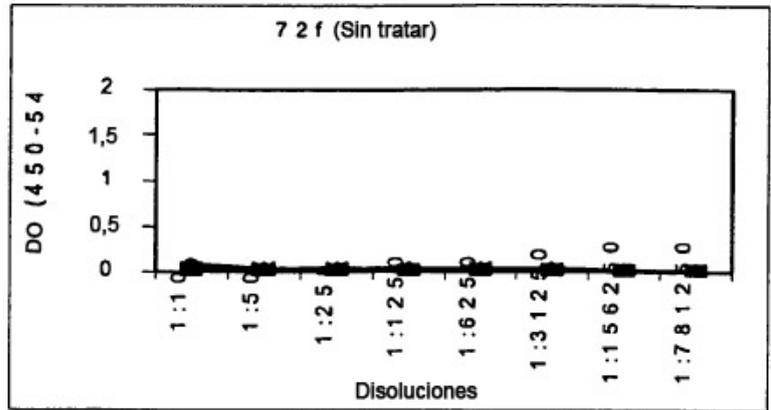


Figura 3

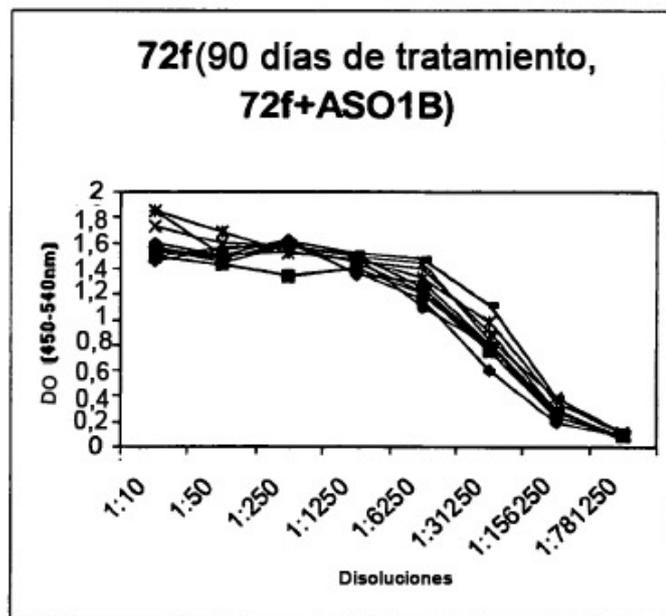
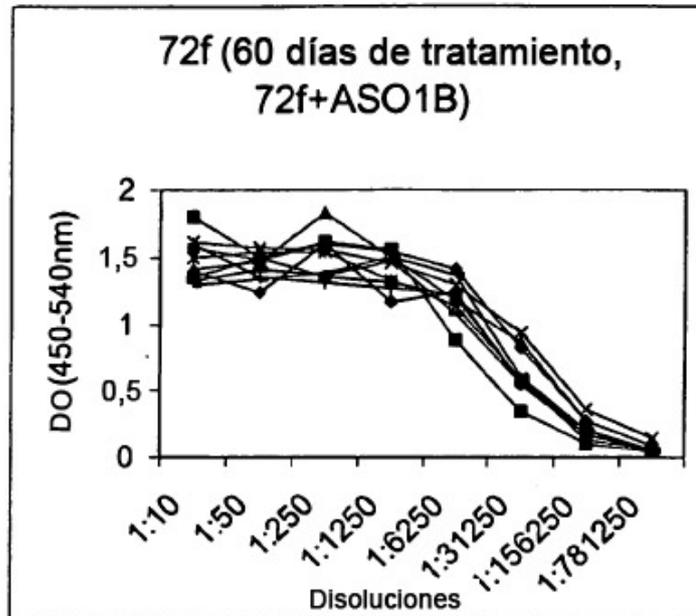


Figura 4

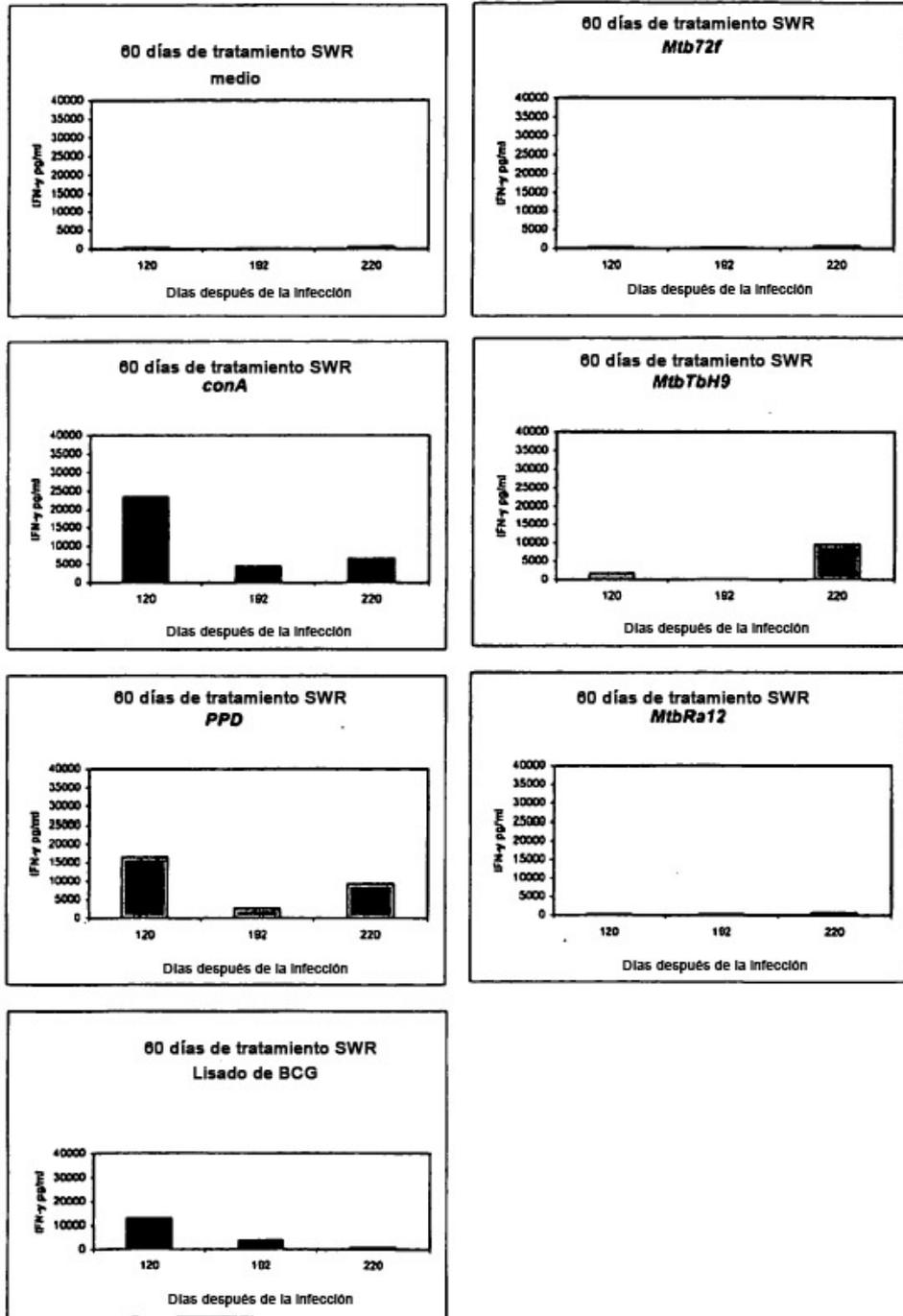


Figura 5

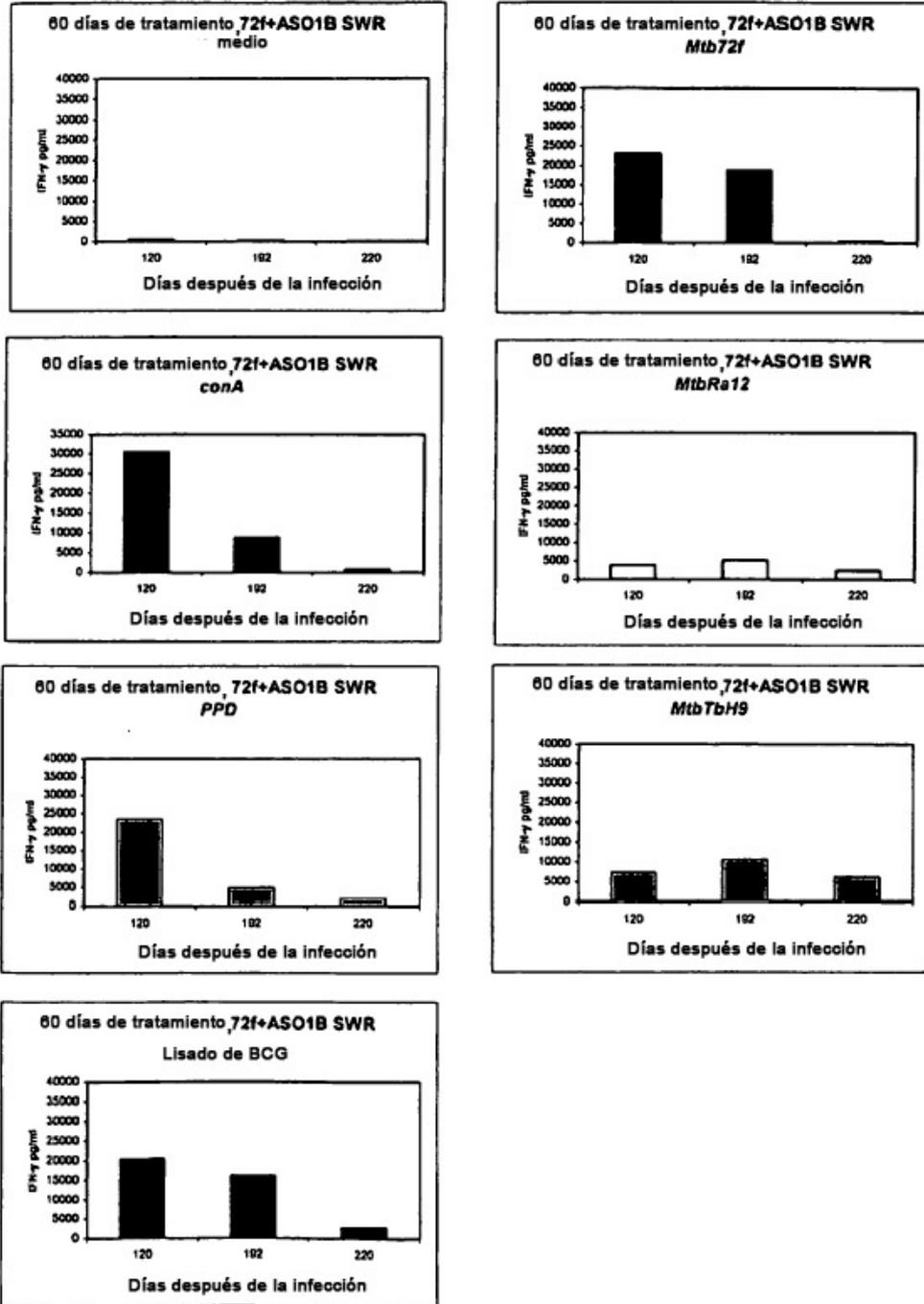
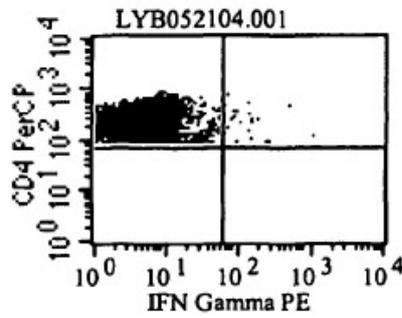
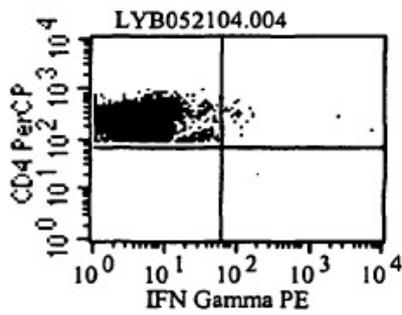


Figura 6A



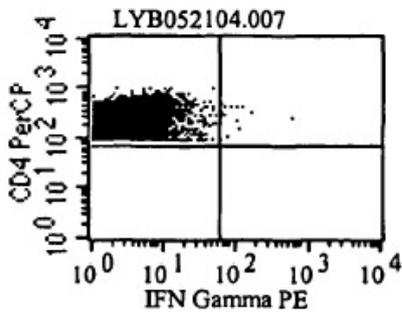
Archivo: LYB052104.001
Ejemplo ID: SWR 1 medio
Puerta: CD4
Casos cerrados: 10129
Casos totales: 236412

Quad	Casos	% cerrado	% Total
UL	10110	99,81	4,28
UR	19	0,19	0,01
LL	0	0,00	0,00
LR	0	0,00	0,00



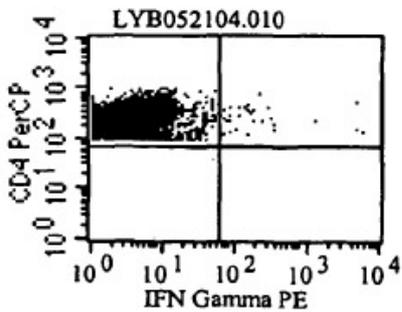
Archivo: LYB052104.004
Ejemplo ID: SWR 2 medio
Puerta: CD4
Casos cerrados: 15138
Casos totales: 184156

Quad	Casos	% cerrado	% Total
UL	15114	99,84	8,21
UR	24	0,16	0,01
LL	0	0,00	0,00
LR	0	0,00	0,00



Archivo: LYB052104.007
Ejemplo ID: SWR 4 medio
Puerta: CD4
Casos cerrados: 19494
Casos totales: 180809

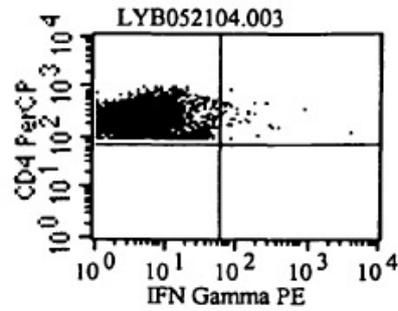
Quad	Casos	% cerrado	% Total
UL	19483	99,94	10,78
UR	11	0,06	0,01
LL	0	0,00	0,00
LR	0	0,00	0,00



Archivo: LYB052104.010
Ejemplo ID: SWR 5 medio
Puerta: CD4
Casos cerrados: 14295
Casos totales: 193256

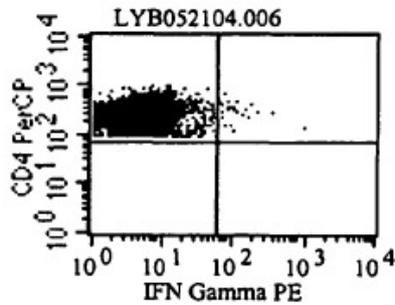
Quad	Casos	% cerrado	% Total
UL	14265	99,79	7,38
UR	30	0,21	0,02
LL	0	0,00	0,00
LR	0	0,00	0,00

Figura 6B



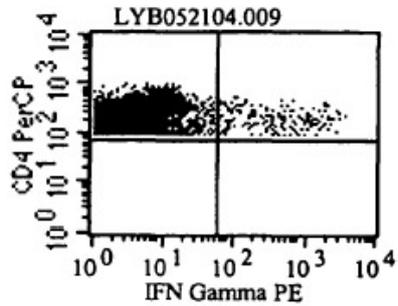
Archivo: LYB052104.003
Ejemplo ID: SWR 1 72F
Puerta: CD4
Casos cerrados: 10267
Casos totales: 232880

Quad	Casos	% cerrado	% Total
UL	10228	99,72	4,39
UR	29	0,28	0,01
LL	0	0,00	0,00
LR	0	0,00	0,00



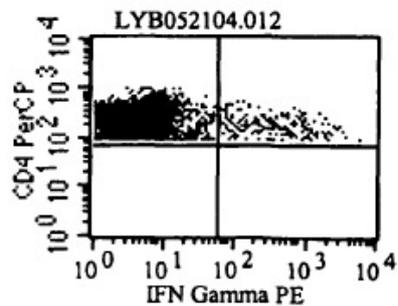
Archivo: LYB052104.006
Ejemplo ID: SWR 2 72F
Puerta: CD4
Casos cerrados: 15634
Casos totales: 190322

Quad	Casos	% cerrado	% Total
UL	15609	99,84	8,20
UR	25	0,16	0,01
LL	0	0,00	0,00
LR	0	0,00	0,00



Archivo: LYB052104.009
Ejemplo ID: SWR 4 72F
Puerta: CD4
Casos cerrados: 19555
Casos totales: 188968

Quad	Casos	% cerrado	% Total
UL	19384	99,13	10,37
UR	171	0,87	0,09
LL	0	0,00	0,00
LR	0	0,00	0,00



Archivo: LYB052104.012
Ejemplo ID: SWR 5 72F
Puerta: CD4
Casos cerrados: 14045
Casos totales: 200028

Quad	Casos	% cerrado	% Total
UL	13849	98,60	6,92
UR	196	1,40	0,10
LL	0	0,00	0,00
LR	0	0,00	0,00

Figura 7

	Grupo 1		Grupo 2		Grupo 3		Grupo 4		Grupo 5	
	Media	72F								
Stims										
RNg+ % cerrado	0,19	0,28	0,16	0,16	0,31	0,8	0,06	0,87	0,21	1,4
% Total	0,01	0,01	0,01	0,01	0,04	0,1	0,01	0,09	0,02	0,1
RNg+ % cerrado	1,78	1,56	0,45	0,56	0,42	2,87	0,34	0,59	0,22	0,61
% Total	0,05	0,04	0,02	0,03	0,03	0,18	0,03	0,05	0,01	0,03

Figura 8

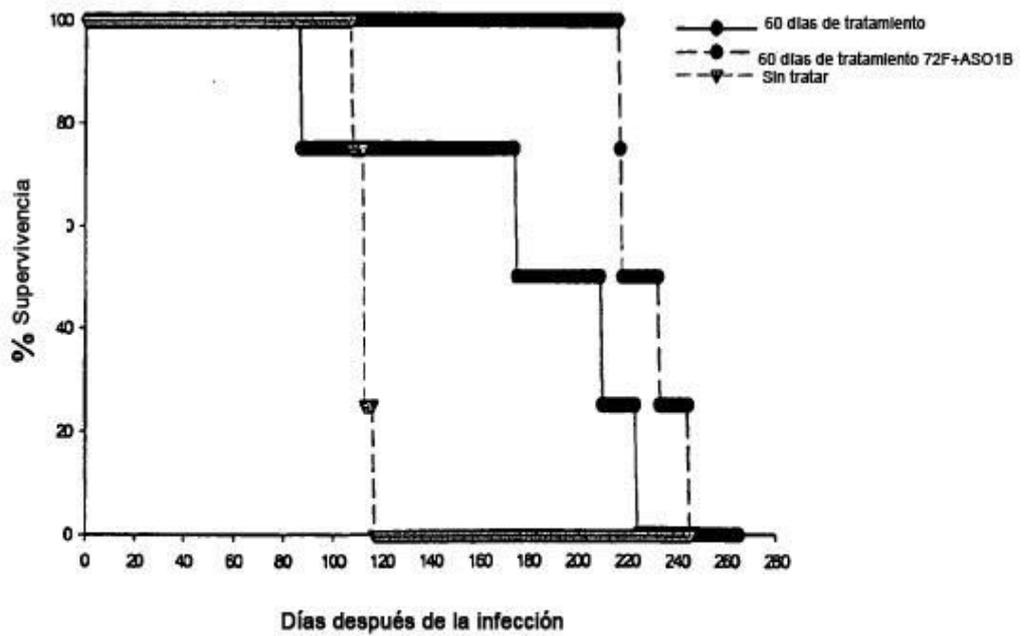


Figura 9

