



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 364 520**

51 Int. Cl.:
C12Q 1/68 (2006.01)
C12N 1/21 (2006.01)
G01N 33/53 (2006.01)
C12N 15/11 (2006.01)
A61K 39/04 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **08015511 .2**
96 Fecha de presentación : **10.08.1999**
97 Número de publicación de la solicitud: **2006395**
97 Fecha de publicación de la solicitud: **24.12.2008**

54 Título: **Marco de lectura abierto Rv2660c de *M. tuberculosis* y su utilización.**

30 Prioridad: **25.08.1998 US 97936 P**
25.05.1999 US 318191

45 Fecha de publicación de la mención BOPI:
06.09.2011

45 Fecha de la publicación del folleto de la patente:
06.09.2011

73 Titular/es: **The Board of Trustees of the Leland
Stanford Junior University**
900 Welch Road, Suite 350
Palo Alto, California 94304-1850, US

72 Inventor/es: **Behr, Marcel;**
Small, Peter;
Schoolnik, Gary y
Wilson, Michael A

74 Agente: **Ponti Sales, Adelaida**

ES 2 364 520 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

Marco de lectura abierto Rv2660c de *M. Tuberculosis* y su utilización

5 [0001] La tuberculosis es una lacra humana antigua que sigue siendo un importante problema de salud pública en todo el mundo. Es una epidemia en curso de proporciones asombrosas. Aproximadamente una de cada tres personas en el mundo está infectada con *Mycobacterium tuberculosis*, y tiene un riesgo de por vida del 10% de avance de la infección a la enfermedad clínica. Aunque la tuberculosis se puede tratar, se estima que 2,9 millones de personas murieron a causa de la enfermedad el año pasado.

10 [0002] Hay problemas significativos con la dependencia en el tratamiento farmacológico para controlar las infecciones activas de *M. tuberculosis*. La mayoría de las regiones con altos índices de infección son los países menos desarrollados, que sufren de una falta de servicios de salud accesibles fácilmente, servicios de diagnóstico y los antibióticos adecuados frente a *M. tuberculosis*. Incluso cuando estos están disponibles, el cumplimiento por parte del paciente suele ser deficiente debido al largo régimen necesario para el tratamiento completo, y las cepas resistentes a múltiples fármacos son cada vez más comunes.

15 [0003] La prevención de la infección eludiría los problemas de tratamiento, por lo que la vacunación contra la tuberculosis se realiza extensamente en las regiones endémicas. Alrededor de 100 millones de personas al año son vacunadas con bacilo de Calmette-Guerin (BCG) vivo. El BCG tiene la gran ventaja de ser barato y puede administrarse con facilidad bajo circunstancias menos que óptimas, con pocas reacciones adversas. Desafortunadamente, la vacuna es muy variable en su eficacia, proporcionando en cualquier lugar de un 0 a un 80% de protección contra la infección con *M. tuberculosis*.

20 [0004] El BCG tiene una historia interesante. Se trata de una cepa atenuada de *M. bovis*, un pariente muy cercano del *M. tuberculosis*. La cepa de *M. bovis* que se convirtió en BCG fue aislada de una vaca en la década de 1800 por parte de un bacteriólogo llamado Nocard, por lo que fue llamado bacilo de Nocard. La atenuación del bacilo Nocard tuvo lugar desde 1908 hasta 1921, a lo largo de 230 pasos *in vitro*. A partir de entonces, se cultiva ampliamente en todo el mundo, resultando en cientos adicionales y a veces miles de pasos *in vitro*. A lo largo de sus muchos años en el laboratorio, ha habido la selección de la reacción cruzada con la prueba cutánea de la tuberculina, y para disminuir los efectos secundarios. El resultado neto ha sido un agente patógeno debilitado considerablemente, lo que puede ser ineficaz en la producción de una respuesta inmunológica adecuada.

25 [0005] Nuevas vacunas contra la tuberculosis se necesitan con urgencia para la población general en las regiones endémicas, para las personas infectadas por el VIH, así como profesionales de la salud que puedan estar expuestos al bacilo de la tuberculosis. Se han desarrollado vacunas de ADN recombinante que llevan genes protectores de *M. tuberculosis* virulenta usando fasmidos de transporte para transferir material genético de una especie micobacteriana a otra, por ejemplo, ver la patente US 5.776.465. Al desarrollo de la vacuna de la tuberculosis se le debe dar una alta prioridad en los actuales objetivos de la investigación médica.

30 *Literatura relevante*

35 [0006] Mahairas et al. (1995) J Bacteriol 178 (5): 1274-1282 proporciona un análisis molecular de las diferencias genéticas entre BCG de *Mycobacterium bovis* y *M. bovis* virulenta. La hibridación genómica sustractiva fue utilizada para identificar las diferencias genéticas entre *M. bovis* virulenta y *M. tuberculosis* y BCG no virulenta, la patente US 5.700.683 se dirige a estas diferencias genéticas.

40 [0007] Cole et al. (1998) Nature 393: 537-544 han descrito el genoma completo de *M. tuberculosis*. Para obtener la secuencia contigua del genoma, se utilizó un enfoque combinado que implicó el análisis de secuencia sistemática de clones seleccionados de gran inserción, así como clones aleatorios de pequeña inserción de una librería de disparo del genoma completo. Esto culminó en una secuencia compuesta de 4.411.529 pares de bases, con un contenido de G + C del 65,6%. 3.924 marcos de lectura abierta se identificaron en el genoma, lo que representa ~91% de la capacidad potencial de codificación.

45 [0008] La secuencia genómica de *Mycobacterium tuberculosis* (*M.tb.*) está disponible en varios sitios de Internet, incluyendo <http://www.cric.com/htdocs/tuberculosis/index.html> y <http://www.sanger.ac.uk/pathogen>.

DESCRIPCIÓN RESUMIDA DE LA INVENCION

50 [0009] La presente invención proporciona una composición de inmunógenos que comprende un polipéptido codificado por como mínimo 25 nucleótidos del marco de lectura abierto Rv2660c, una micobacteria alterada genéticamente que comprende dicho marco de lectura abierto y métodos para la utilización de los productos de dicho marco de lectura abierto.

55 [0010] Se proporcionan marcadores genéticos que distinguen entre las cepas del *Mycobacterium tuberculosis* complejo, en particular entre cepas no virulentas y virulentas. Las cepas de interés incluyen *M.bovis*, cepas de BCG de *M.bovis*,

5 *M.tuberculosis* (*M.tb.*) aislado, y bacteriófagos que infectan las micobacterias. Los marcadores genéticos se utilizan para los ensayos, por ejemplo, inmunoensayos, que distinguen entre las cepas, tal como para diferenciar entre la inmunización de BCG y la infección *M.tb.* Los productos de proteína pueden producirse y utilizarse como inmunógenos, en el cribado de fármacos, etc. Los marcadores son útiles en la construcción modificada genéticamente de células *M.tb.* o *M.bovis* que tienen características mejoradas de la vacuna.

DESCRIPCIÓN DETALLADA DE LAS REALIZACIONES

1.0 [0011] Se identifican supresiones genéticas específicas, y específicamente Rv2654c, que sirven como marcadores para distinguir entre cepas de micobacterias no virulentas y virulentas, incluyendo cepas de *M.bovis*, BCG de *M.bovis*, *M.tuberculosis* (*M.tb.*) aislado y bacteriófagos que infectan las micobacterias. Esta supresión se utiliza como marcador genético para distinguir entre las micobacterias diferentes. La supresión puede introducirse en *M.tb.* o *M.bovis* mediante procedimientos recombinantes con el fin de producir una cepa patógena no virulenta. Por otra parte, el gen suprimido se
1.5 identifica en la secuencia del genoma *M.tb.*, y luego de nuevo mediante procedimientos recombinantes en BCG u otras cepas de vacuna, para mejorar la eficacia de la vacunación.

2.0 [0012] La supresión de la invención se identifica mediante hibridaciones de ADN comparativo de secuencias genómicas de micobacterias a una micromatriz de ADN que comprende las secuencias representativas de las secuencias codificadoras de *M.tb.*. Las supresiones se asignan entonces a la secuencia del genoma *M.tb.* conocida para identificar específicamente el gen o genes suprimidos, y para caracterizar la secuencia de nucleótidos de la región suprimida.

2.5 [0013] Los ácidos nucleicos que comprenden la supresión y las uniones proporcionadas se utilizan en una variedad de aplicaciones. Se pueden obtener sondas de hibridación a partir de la secuencia *M. tb.* conocida que se corresponde a las secuencias suprimidas. Estas sondas son útiles para distinguir entre las micobacterias. Por ejemplo, hay una probabilidad del 10% de que una persona infectada con *M. tb.* progrese a una enfermedad clínica, pero esa probabilidad puede variar dependiendo de la cepa infectante particular. Se utiliza un análisis para detectar la presencia o la ausencia de las supresiones previstas a continuación como "*M. tb* viable" para distinguir entre diferentes cepas de *M. tb.* Las supresiones son también útiles para identificar si un paciente que es positivo para la prueba cutánea de la tuberculina ha
3.0 sido infectado con *M. tb* o con BCG.

3.5 [0014] En otra realización, las micobacterias son alteradas genéticamente para suprimir las secuencias identificadas aquí como ausentes en las cepas atenuadas, pero presentes en las cepas patógenas, por ejemplo, supresiones en BCG, pero presentes en H37Rv de *M. tb.* Estas cepas modificadas genéticamente pueden proporcionar vacunas superiores a los actuales presentes aislamientos de BCG en uso. Alternativamente, las cepas de BCG puede ser "reconstruidas" para parecerse más a *M. tb* de tipo salvaje mediante la inserción de nuevo de algunas de las secuencias suprimidas en el genoma. Como los productos de proteína de las secuencias suprimidas se expresan en especies de micobacterias virulentas, las proteínas codificadas son útiles como inmunógenos para la vacunación.

4.0 [0015] La atenuación (pérdida de virulencia) en BCG se atribuye a la pérdida de material genético en una serie de sitios en todo el genoma. La selección a lo largo del tiempo para menores efectos secundarios resultantes de la inmunización con BCG, mientras que conserva la reactividad cruzada con la prueba cutánea de la tuberculina, ha proporcionado un excelente cribado para las secuencias que generan efectos secundarios. La identificación de las supresiones que varían entre los aislados de BCG que identifican estas secuencias, que se pueden utilizar en el cribado de fármacos y el análisis biológico para el papel de los genes suprimidos como causa de efectos secundarios adversos y patogenicidad.
4.5

Identificación de Marcadores de Supresión Complejos de *M. Tuberculosis*

5.0 [0016] Se proporcionan secuencias de ácido nucleico que son marcadores para micobacterias específicas, incluyendo *M. tb.*, *M. bovis*, BCG y bacteriófagos. Las supresiones aparecen en la Tabla 1. La ausencia o la presencia de estas secuencias de marcadores son características del aislamiento indicado, o cepa. Por tanto, proporcionan una característica única para la identificación de las micobacterias indicadas. Las supresiones se identifican por su marco de lectura abierto *M. tb.* (nomenclatura "RV") que corresponde a una secuencia genética conocida, y puede ser accesible tal como se ha citado previamente. Las uniones de las supresiones se proporcionan mediante la designación de la posición en la secuencia *M. tb.* disponible públicamente.
5.5

Tabla 1

Rd.	Rv_num	orf_id	Punto de Ruptura
RD01	Rv3871	MTV027.06	"H37Rv, segmento 160: 7534, 16989"
RD01	Rv3872	MTV027.07	"H37Rv, segmento 160: 7534, 16989"
RD01	Rv3873	MTV027.08	"H37Rv, segmento 160: 7534, 16989"
RD01	Rv3874	MTV027.09	"H37Rv, segmento 160: 7534, 16989"

ES 2 364 520 T3

Rd.	Rv_num	orf_id	Punto de Ruptura
RD01	Rv3875	MTV027.10	"H37Rv, segmento 160: 7534, 16989"
RD01	Rv3876	MTV027.11	"H37Rv, segmento 160: 7534, 16989"
RD01	Rv3877	MTV027.12	"H37Rv, segmento 160: 7534, 16989"
RD01	Rv3878	MTV027.13	"H37Rv, segmento 160: 7534, 16989"
RD01	Rv3879c	MTV027.14c	"H37Rv, segmento 160: 7534, 16989"
RD02	Rv1988	MTCY39.31c	"Segmento H37Rv 88: 14211, el segmento de 89: 8598"
RD02	Rv1987	MTCY39.32c	"Segmento H37Rv 88: 14211, el segmento de 89: 8598"
RD02	Rv1986	MTCY39.33c	"Segmento H37Rv 88: 14211, el segmento de 89: 8598"
RD02	Rv1985c	MTCY39.34	"Segmento H37Rv 88: 14211, el segmento de 89: 8598"
RD02	Rv1984c	MTCY39.35	"Segmento H37Rv 88: 14211, el segmento de 89: 8598"
RD02	Rv1983	MTCY39.36c	"Segmento de 88 H37Rv: 14211, el segmento de 89: 8598"
RD02	Rv1982c	MTCY39.37	"Segmento H37Rv 88: 14211, el segmento de 89: 8598"
RD02	Rv1981c	MTCY39.38	"Segmento H37Rv 88: 14211, el segmento de 89: 8598"
RD02	Rv1980c	MTCY39.39	"Segmento H37Rv 88: 14211, el segmento de 89: 8598"
RD02	Rv1979c	MTCY39.40	"Segmento H37Rv 88: 14211, el segmento de 89: 8598"
RD02	Rv1978	MTV051.16	"Segmento H37Rv 88: 14211, el segmento de 89: 8598"
RD03	Rv1586c	MTCY336.18	"H37Rv, segmento 70: 7677, 16923"
RD03	Rv1585c	MTCY336.19	"H37Rv, segmento 70: 7677, 16923"
RD03	Rv1584c	MTCY336.20	"H37Rv, segmento 70: 7677, 16923"
RD03	Rv1583c	MTCY336.21	"H37Rv, segmento 70: 7677, 16923"
RD03	Rv1582c	MTCY336.22	"H37Rv, segmento 70: 7677, 16923"
RD03	Rv1581c	MTCY336.23	"H37Rv, segmento 70: 7677, 16923"
RD03	Rv1580c	MTCY336.24	"H37Rv, segmento 70: 7677, 16923"
RD03	Rv1579c	MTCY336.25	"H37Rv, segmento 70: 7677, 16923"
RD03	Rv1578c	MTCY336.26	"H37Rv, segmento 70: 7677, 16923"
RD03	Rv1577c	MTCY336.27	"H37Rv, segmento 70: 7677, 16923"
RD03	Rv1576c	MTCY336.28	"H37Rv, segmento 70: 7677, 16923"
RD03	Rv1575	MTCY336.29c	"H37Rv, segmento 70: 7677, 16923"
RD03	Rv1574	MTCY336.30c	"H37Rv, segmento 70: 7677, 16923"
RD03	Rv1573	C MTCY336.31	"H37Rv, segmento 70: 7677, 16923"
RD04	Rv0221	MTCY08D5.16	"H37Rv, segmento 12: 17432,19335"
RD04	Rv0222	MTCY08D5.17	"H37Rv, segmento 12: 17432,19335"
RD04	Rv0223c	MTCY08D5.18	"H37Rv, segmento 12: 17432,19335"
RD05	Rv3117	MTCY164.27	"H37Rv, segmento 135: 27437,30212"
RD05	Rv3118	MTCY164.28	"H37Rv, segmento 135: 27437,30212"
RD05	Rv3119	MTCY164.29	"H37Rv, segmento 135: 27437,30212"
RD05	Rv3120	MTCY164.30	"H37Rv, segmento 135: 27437,30212"
RD05	Rv3121	MTCY164.31	"H37Rv, segmento 135: 27437,30212"
RD06	Rv1506c	MTCY277.28c	"H37Rv, segmento 65: 23614, 36347"
RD06	Rv1507c	MTCY277.29c	"H37Rv, segmento 65: 23614, 36347"
RD06	Rv1508c	MTCY277.30c	"H37Rv, segmento 65: 23614, 36347"
RD06	Rv1509	MTCY277.31	"H37Rv, segmento 65: 23614, 36347"

ES 2 364 520 T3

Rd.	Rv_num	orf_id	Punto de Ruptura
RD06	Rv1510	MTCY277.32	"H37Rv, segmento 65: 23614, 36347"
RD06	Rv1511	MTCY277.33	"H37Rv, segmento 65: 23614, 36347"
RD06	Rv1512	MTCY277.34	"H37Rv, segmento 65: 23614, 36347"
RD06	Rv1513	MTCY277.35	"H37Rv, segmento 65: 23614, 36347"
RD06	Rv1514c	MTCY277.36c	"H37Rv, segmento 65: 23614, 36347"
RD06	Rv1515c	MTCY277.37c	"H37Rv, segmento 65: 23614, 36347"
RD06	Rv1516c	MTCY277.38c	"H37Rv, segmento 65: 23614, 36347"
RD07	Rv2346c	MTCY98.15c	"H37Rv, segmento 103: 17622, 26584"
RD07	Rv2347c	MTCY98.16c	"H37Rv, segmento 103: 17622, 26584"
RD07	Rv2348c	MTCY98.17c	"H37Rv, segmento 103: 17622, 26584"
RD07	Rv2349c	MTCY98.18c	"H37Rv, segmento 103: 17622, 26584"
RD07	Rv2350c	MTCY98.19c	"H37Rv, segmento 103: 17622, 26584"
RD07	Rv2351c	MTCY98.20c	"H37Rv, segmento 103: 17622, 26584"
RD07	Rv2352c	MTCY98.21c	"H37Rv, segmento 103: 17622, 26584"
RD07	Rv2353c	MTCY98.22c	"H37Rv, segmento 103: 17622, 26584"
RD08	Rv0309	MTCY63.14	"H37Rv, segmento de 16: 17018, 20446"
RD08	Rv0310c	MTCY63.15c	"H37Rv, segmento de 16: 17018, 20446"
RD08	Rv0311	MTCY63.16	"H37Rv, segmento de 16: 17018, 20446"
RD08	Rv0312	MTCY63.17	"H37Rv, segmento de 16: 17018, 20446"
RD09	Rv3623	MTCY15C10.29c	"H37Rv, segmento 153: 21131, el segmento de 154: 2832"
RD09	Rv3622c	MTCY15C10.30	"H37Rv, segmento 153: 21131, el segmento de 154: 2832"
RD09	Rv3621c	MTCY15C10.31	"H37Rv, segmento de 15,3: 21131, el segmento de 154: 2832"
RD09	Rv3620c	MTCY15C10.32	"H37Rv, segmento 153: 21131, el segmento de 154: 2832"
RD09	Rv3619c	MTCY15C10.33	"H37Rv, segmento 153: 21131, el segmento de 154: 2832"
RD09	Rv3618	MTCY15C10.34c	"H37Rv, segmento 153: 21131, el segmento de 154: 2832"
RD09	Rv3617	MTCY15C10.35c	"H37Rv, segmento 153: 21131, el segmento de 154: 2832"
RD10	Rv1257c	MTCY50.25	"Segmento H37Rv 55: 3689, 6696"
RD10	Rv1256c	MTCY50.26	"Segmento H37Rv 55: 3689, 6696"
RD10	Rv1255c	MTCY50.27	"Segmento H37Rv 55: 3689, 6696"
RD11	Rv3429	MTCY77.01	"H37Rv, segmento 145: 30303 para segmentar 146: 1475"
RD11	Rv3428c	MTCY78.01	"H37Rv, segmento 145: 30303 para segmentar 146: 1475"
RD11	Rv3427c	MTCY78.02	"H37Rv, segmento 145: 30303 para segmentar 146: 1475"
RD11	Rv3426	MTCY78.03c	"H37Rv, segmento 145: 30303 para segmentar 146: 1475"
RD11	Rv3425	MTCY78.04c	"H37Rv, segmento 145: 30303 para segmentar 146: 1475"
RD12	Rv2072c	MTCY49.11c	"Segmento H37Rv 93: 9301, 11331"
RD12	Rv2073c	MTCY49.12c	"Segmento H37Rv 93: 9301, 11331"
RD12	Rv2074	MTCY49.13	"Segmento H37Rv 93: 9301, 11331"
RD12	Rv2075c	MTCY49.14c	"Segmento H37Rv 93: 9301, 11331"
RD13bis	Rv2645	MTCY441.15	"H37Rv, segmento 118: 12475, 23455"
RD13bis	Rv2646	MTCY441.16	"H37Rv, segmento 118: 12475,23455"
RD13bis	Rv2647	MTCY441.17	"H37Rv, segmento 118: 12475, 23455"
RD13bis	Rv2648	MTCY441.17A	"H37Rv, segmento 118: 12475, 23455"

ES 2 364 520 T3

Rd.	Rv_num	orf_id	Punto de Ruptura
RD13bis	Rv2649	MTCY441.18	"H37Rv, segmento 118: 12475, 23455"
RD13bis	Rv2650c	MTCY441.19	"H37Rv, segmento 118: 12475, 23455"
RD13bis	Rv2651c	MTCY441.20c	"H37Rv, segmento 118: 12475, 23455"
RD13bis	Rv2652c	MTCY441.21c	"H37Rv, segmento 118: 12475, 23455"
RD13bis	Rv2653c	MTCY441.22c	"H37Rv, segmento 118: 12475, 23455"
RD13bis	Rv2654c	MTCY441.23c	"H37Rv, segmento 118: 12475, 23455"
RD13bis	Rv2655c	MTCY441.24c	"H37Rv, segmento 118: 12475, 23455"
RD13bis	Rv2656c	MTCY441.25c	"H37Rv, segmento 118: 12475, 23455"
RD13bis	Rv2657c	MTCY441.26c	"H37Rv, segmento 118: 12475,23455"
RD13bis	Rv2658c	MTCY441.27c	"H37Rv, segmento 118: 12475,23455"
RD13bis	Rv2659c	MTCY441.28c	"H37Rv, segmento 118: 12475,23455"
RD13bis	Rv2660c	MTCY441.29c	"H37Rv, segmento 118: 12475, 23455"
RD14	Rv1766	MTCY28.32	"Segmento H37Rv 79: 30573, 39642"
RD14	Rv1767	MTCY28.33	"Segmento H37Rv 79: 30573, 39642"
RD14	Rv1768	MTCY28.34	"Segmento H37Rv 79: 30573, 39642"
RD14	Rv1769	MTCY28.35	"Segmento H37Rv 79: 30573, 39642"
RD14	Rv1770	MTCY28.36	"Segmento H37Rv 79: 30573, 39642"
RD14	Rv1771	MTCY28.37	"Segmento H37Rv 79: 30573, 39642"
RD14	Rv1772	MTCY28.38	"Segmento H37Rv 79: 30573, 39642"
RD14	Rv1773c	MTCY28.39	"Segmento H37Rv 79: 30573, 39642"
RD15	Rv1963c	MTV051.01c	"Segmento H37Rv 88: 1153, 13873"
RD15	Rv1964	MTV051.02	"Segmento H37Rv 88: 1153, 13873"
RD15	Rv1965	MTV051.03	"Segmento H37Rv 88: 1153, 13873"
RD15	Rv1966	MTV051.04	"Segmento H37Rv 88: 1153, 13873"
RD15	Rv1967	MTV051.05	"Segmento H37Rv 88: 1153, 13873"
RD15	Rv1968	MTV051.06	"Segmento H37Rv 88: 1153, 13873"
RD15	Rv1969	MTV051.07	"Segmento H37Rv 88: 1153, 13873"
RD15	Rv1970	MTV051.08	"Segmento H37Rv 88: 1153, 13873"
RD15	Rv1971	MTV051.09	"Segmento H37Rv 88: 1153, 13873"
RD15	Rv1972	MTV051.10	"Segmento H37Rv 88: 1153,13873"
RD15	Rv1973	MTV051.11	"Segmento H37Rv 88: 1153, 13873"
RD15	Rv1974	MTV051.12	"Segmento H37Rv 88: 1153, 13873"
RD15	Rv1975	MTV051.13	"Segmento H37Rv 88: 1153, 13873"
RD15	Rv1976c	MTV051.14	"Segmento H37Rv 88: 1153, 13873"
RD15	Rv1977	MTV051.15	"Segmento H37Rv 88: 1153, 13873"
RD16	Rv3405c	MTCY78.23	"H37Rv, segmento 145: 5012, 12621"
RD16	Rv3404c	MTCY78.24	"H37Rv, segmento 145: 5012, 12621"
RD16	Rv3403c	MTCY78.25	"H37Rv, segmento 145: 5012, 12621"
RD16	Rv3402c	MTCY78.26	"H37Rv, segmento 145: 5012, 12621"
RD16	Rv3401	MTCY78.27c	"H37Rv, segmento 145: 5012, 12621"
RD16	Rv3400	MTCY78.28c	"H37Rv, segmento 145: 5012, 12621"

La columna "Rv" indica secuencia *M. tb* pública, marco de lectura abierto. Las cepas de BCG se obtuvieron de la siguiente manera:

Tabla 2. Cepas empleadas en el estudio de filogenia BCG

5

Nombre de la cepa	Sinónimo	Fuente	Descriptores
BCG-Rusia	Moscú	ATCC	# 35740
BCG-Moreau	Brasil	ATCC	# 35736
BCG-Moreau	Brasil	IAF	1958 de fecha
BCG-Moreau	Brasil	IAF	fechado en 1961
BCG-Japón	Tokio	ATCC	# 35737
BCG-Japón	Tokio	IAF	fechado en 1961
BCG-Japón	Tokio	JATA	cepa de vacuna
BCG-Japón	Tokio	JATA	cepa cáncer de vejiga
BCG-Japón	Tokio	JATA	aislado clínico-adenitis
BCG-Suecia	Gotemburgo	ATCC	# 35732
BCG-Suecia	Gotemburgo	IAF	fechado en 1958
BCG-Suecia	Gotemburgo	SSI	lote de producción, Copenhague
BCG Phipps	Filadelfia	ATCC	# 35744
BCG-Dinamarca	Danesa 1331	ATCC	# 35733
BCG-Copenhague		ATCC	# 27290
BCG-Copenhague		IAF	fechado en 1961
BCG-Tice	Chicago	Vacuna	fechado en 1973
BCG-Tice	Chicago	ATCC	# 35743
BCG-Frappier	Montreal	IAF	lote primario, 1973
BCG-Frappier, INH-resistente	Montreal-R	IAF	lote primario, 1973
BCG-Frappier	Montreal	IAF	pasaje 946
BCG-Connaught	Toronto	CL	tratamiento del cáncer de vejiga
BCG-Birkhaug		ATCC	# 35731
BCG-Praga	Checo	SSI	liofilizado 1968
BCG-Glaxo		Vacuna	fechado en 1973
BCG-Glaxo		ATCC	# 35741
BCG-Pasteur		IAF	pasaje 888
BCG-Pasteur		IAF	fechado en 1961
BCG-Pasteur		IP	1173P2-B
BCG-Pasteur		IP	1173P2-C
BCG-Pasteur		IP	asilado clínico # 1
BCG-Pasteur		IP	aislado clínico # 2
BCG-Pasteur		ATCC	# 35734

Abreviaciones: IP = Institut Pasteur, París, Francia; IAF = Institut Armand-Frappier, Laval, Canadá, ATCC = American Type Culture Collection, Rockville, Maryland, EE.UU., SSI = Statens Serum Institute, Copenhague, Dinamarca; CL = Laboratorios Connaught, Willowdale, Canadá, JATA = Asociación AntiTuberculosis de Japón; INH = isoniazida. BCG canadiense se refiere a BCG-Montreal y BCG-Toronto, derivándose éstos de los primeros.

[0017] En la realización del procedimiento de cribado inicial, el ADN genómico se aisló a partir de dos cultivos de células microbianas de micobacterias. Las dos preparaciones de ADN están etiquetadas, si se utiliza una etiqueta diferente para el primer y segundo cultivos microbianos, típicamente utilizando nucleótidos conjugados a un fluorocromo que emite a una longitud de onda sustancialmente diferente de la de los nucleótidos marcados con fluorocromo utilizados para etiquetar la sonda seleccionada. Las cepas utilizadas fueron la cepa de referencia de *Mycobacterium tuberculosis* (H37Rv), otras cepas *M. tb.* de laboratorio, tales como H37Ra, la cepa O, *M. tb.* aislados clínicos, la cepa de referencia de *Mycobacterium bovis*, y diferentes cepas de BCG de *Mycobacterium bovis*.

[0018] Las dos preparaciones de ADN se mezclan, y la hibridación competitiva se lleva a cabo en una micromatriz que representa todos los marcos de lectura abierta en el genoma del microbio de prueba, usualmente H37Rv. La hibridación de las secuencias marcadas se realiza de acuerdo con procedimientos bien conocidos en la técnica. En una realización preferida, las dos sondas se combinan para proporcionar una hibridación competitiva en una sola micromatriz. La hibridación puede realizarse bajo condiciones que varían en cuanto al rigor, preferentemente bajo condiciones de alta astringencia (*por ejemplo*, 4x SSC, 10% SDS, 65°C) para permitir la hibridación de secuencias complementarias que tienen una homología extensiva (*por ejemplo*, que tienen al menos un 85% de identidad de secuencia, preferentemente al menos un 90% de identidad de secuencia, más preferiblemente con al menos un 95% de identidad de secuencia). Si las secuencias objetivo son secuencias nativas, la hibridación se realiza preferentemente en condiciones que permiten la hibridación de sólo secuencias muy homólogas (*por ejemplo* al menos un 95% a un 100% de identidad de secuencia).

[0019] Hibridación fluorescente de dos colores se utiliza para ensayos de la representación de la librería no seleccionada en relación con la librería seleccionada (*es decir*, para detectar la hibridación de la sonda no seleccionada respecto a la sonda seleccionada). De la relación de un color al otro, para cualquier elemento de la matriz particular, se puede determinar la abundancia relativa de esa secuencia en la librería seleccionada y no seleccionada. Además, la comparación de la hibridación de las sondas seleccionadas y no seleccionadas ofrece un control interno para el ensayo. La ausencia de señal de la cepa de referencia, en comparación con H37Rv, es indicativa de que el marco de lectura abierto está suprimido en la cepa de prueba. La supresión también se puede mapear mediante análisis Southern blot, y mediante la secuenciación de las regiones que flanquean la supresión.

[0020] Las micromatrices se pueden escanear para detectar la hibridación de las secuencias seleccionadas y no seleccionadas utilizando un microscopio de escaneado láser de construcción habitual tal como se describe en Shalon et al., Res. Genoma. 6:639 (1996). Una exploración por separado, utilizando la línea de excitación adecuada, se realiza para cada uno de los dos fluoróforos utilizados. Las imágenes digitales generadas a partir de la exploración se combinan para su posterior análisis. Para cualquier elemento de la matriz particular, la relación entre la señal fluorescente del ADN de la población celular seleccionada amplificado asociado se compara con la señal fluorescente del ADN de la población celular no seleccionada, y se determina la abundancia relativa de esa secuencia en la librería seleccionado y no seleccionada.

Composición de ácidos nucleicos

[0021] Tal como se utiliza aquí, el término "marcador de supresión" o "marcador" se utiliza para referirse a las secuencias de genomas complejos de *M. tuberculosis* que se eliminan en una o más de las cepas o especies, tal como se indica en la Tabla 1, y específicamente Rv2660cc. Las bacterias del complejo de *M. tuberculosis* incluyen *M. tuberculosis*, *M. bovis*, y BCG, incluido de aislados y cepas variadas en cada especie. Los ácidos nucleicos de interés incluyen la totalidad o una parte de la región suprimida, particularmente marcos de lectura abiertos completos, cebadores de hibridación, regiones promotoras, etc.

[0022] El término "unión" o "unión de supresión" se utiliza para referirse a ácidos nucleicos que comprenden las regiones en la secuencia 3' y 5' que flanquean inmediatamente la supresión. Estas secuencias de unión se utilizan preferentemente como cebadores cortos, *por ejemplo* de aproximadamente 15 nt a aproximadamente 30 nt, que específicamente se hibridan con la unión, pero no a un ácido nucleico que comprende la secuencia genómica no suprimida. Por ejemplo, la supresión encontrada en *M. bovis*, en Rv0221, se corresponde a la secuencia de nucleótidos del genoma de *M. tuberculosis* H37Rv, segmento 12: 17432,19335. La unión comprende las regiones situadas antes de la posición 17342, y después de la posición 19335, *por ejemplo* un ácido nucleico de 20 nucleótidos que comprende la secuencia de H37Rv 17332-17342 unida a 19335-19345.

[0023] Típicamente, estos ácidos nucleicos que comprenden una unión incluirán por lo menos aproximadamente 7 nucleótidos de cada región de flaqueo, es decir, desde la secuencia 3' y desde la secuencia 5' adyacente a la supresión, y puede ser de aproximadamente 10 nucleótidos desde cada región de flaqueo, hasta aproximadamente 15 nucleótidos, o más. Los cebadores de amplificación que hibridan con la secuencia de unión, a la secuencia suprimida, y a las regiones no suprimidas que la flanquean tienen una variedad de usos, tal como se detalla a continuación.

[0024] Las composiciones de ácido nucleico aquí codifican todos o una parte de los marcadores de supresión. Se pueden obtener fragmentos de la secuencia de ADN mediante síntesis de oligonucleótidos químicamente, de acuerdo con procedimientos convencionales, mediante digestión con enzimas de restricción, mediante amplificación por PCR, etc. En su mayor parte, los fragmentos de ADN tendrán por lo menos 25 nt de longitud aproximadamente, usualmente por lo menos 30 nt aproximadamente, más usualmente por lo menos 50 nt aproximadamente. Para su uso en

reacciones de amplificación, tal como PCR, se utilizarán un par de cebadores. La composición exacta de las secuencias de los cebadores no es crítica para la invención, pero para la mayoría de las aplicaciones los cebadores se hibridarán con la secuencia sujeto bajo condiciones estrictas, tal como se conoce en la técnica. Es preferible elegir un par de cebadores que generen un producto de amplificación de al menos unos 50 nt, preferentemente al menos unos 100 nt.

5 Los algoritmos para la selección de las secuencias de los cebadores son generalmente conocidos y están disponibles en paquetes de software comerciales. Los cebadores de amplificación se hibridan, en cadenas complementarias de ADN, y se ceban entre sí.

10 [0025] Usualmente, el ADN se obtendrá substancialmente libre de otras secuencias de ácido nucleico que no incluyen una secuencia marcadora de supresión o fragmento de la misma, generalmente siendo al menos del 50%, usualmente por lo menos un 90% puro aproximadamente y son típicamente "recombinantes", es decir, flanqueado por uno o más nucleótidos con los que normalmente no está asociados en un cromosoma de origen natural.

15 [0026] Para propósitos de cribado, las sondas de hibridación de una o más de las secuencias de supresión pueden utilizarse en reacciones separadas o espacialmente separadas en una matriz de fase sólida, o etiquetados de tal manera que se pueden distinguir entre sí. Los ensayos pueden utilizar ácidos nucleicos que hibridan una o más de las supresiones descritas.

20 [0027] Una matriz puede incluir la totalidad o un subconjunto de los marcadores de supresión en la Tabla 1. Usualmente, esta matriz incluirá al menos dos secuencias diferentes de marcadores de supresión, es decir, supresiones situadas en posiciones únicas en el lugar, y puede incluir todos los marcadores de supresión previstos. Las matrices de interés pueden comprender además otras secuencias genéticas, particularmente otras secuencias de interés para el cribado de la tuberculosis. La secuencia de oligonucleótidos de la matriz será usualmente por lo menos de 12 nt de longitud aproximadamente, puede tener la longitud de las secuencias de marcador de supresión previstas, o puede extenderse a las regiones de flaqueo para generar fragmentos de 100 a 200 nt de longitud. Para ejemplos de matrices, ver Ramsay (1998) Nat. Biotech. 16:40-44; Hacia et al. (1996) 14:441-447 Nature Genetics; Lockhart et al. (1996) Nature Biotechnol. 14:1675-1680, y De Risi et al. (1996) 14:457-460 Nature Genetics.

30 [0028] Los ácidos nucleicos pueden ser de origen natural, *por ejemplo* ADN o ARN, o pueden ser análogos sintéticos, tal como se conoce en la técnica. Estos análogos pueden ser preferibles para su uso como sondas debido a la estabilidad superior bajo condiciones de ensayo. Las modificaciones en la estructura nativa, incluidas las modificaciones en el esqueleto, azúcares o bases heterocíclicas, se han demostrado que aumenta la estabilidad intracelular y la afinidad de unión. Entre las modificaciones útiles en la química del esqueleto hay fosforotioatos; fosforoditioatos, donde tanto los oxígenos no puente son sustituidos con azufre; fosforoamiditos; alquil fosfortriésteres y boranofosfatos. Los derivados de fosfato aquirales incluyen 3'-O'-5'-S-fosforotioato, 3'-S-5'-O-fosforotioato, 3'-5'-CH₂-O-fosfonato y 3'-NH-5'-O-fosforoamidato. Ácidos nucleicos péptidos sustituyen todo el esqueleto de ribosa fosfodiéster con un enlace de péptido.

35 [0029] También se utilizan modificaciones de azúcar para mejorar la estabilidad y la afinidad. El O-anómero de desoxirribosa puede utilizarse, cuando la base se invierte respecto a los b-anómeros naturales. El 2'-OH del azúcar de ribosa puede modificarse para formar 2'-O-metilo o azúcares 2'-O-alilo, que proporcionan resistencia a la degradación sin comprometer la afinidad.

40 [0030] La modificación de las bases heterocíclicas debe mantener el emparejado adecuado de la base. Algunas sustituciones útiles incluyen desoxiuridina para desoxitimidina; 5-etil-2-desoxicidina y 5-bromo-2-desoxicidina para desoxicidina. 5-propinil-2'-desoxiuridina y 5-propinil-2'-desoxicidina han demostrado que aumentan la afinidad y la actividad biológica cuando sustituyen a la desoxitimidina y la desoxicidina, respectivamente.

Composiciones de Polipéptidos

50 [0031] Los marcadores de supresión específicos en la Tabla 1 y en particular Rv2660c se corresponden a marcos de lectura abierta del genoma de *M. tb*, y por lo tanto codifican un polipéptido. Los marcadores objeto pueden utilizarse para la síntesis de una proteína completa, o fragmentos de polipéptido de las mismas, en particular fragmentos correspondientes a dominios funcionales, sitios de unión; *etc.*; e incluyendo fusiones de los polipéptidos objeto a otras proteínas o partes de las mismas. Para la expresión, se puede utilizar un cassette de expresión, que proporciona una región de iniciación transcripcional y translacional, que puede ser inducible o constitutiva, donde la región de codificación está operativa unida bajo el control transcripcional de la región de iniciación de la transcripción, y una región de terminación transcripcional y translacional. Pueden utilizarse varias regiones de iniciación transcripcionales que son funcionales en el huésped de expresión.

60 [0032] Los polipéptidos pueden expresarse en procariontes o eucariotes, de acuerdo con las formas convencionales, dependiendo del propósito de la expresión. Para la producción a gran escala de la proteína, un organismo unicelular, tal como *E. coli*, *B. subtilis*, *S. cerevisiae*, o células de un organismo superior, tal como los vertebrados, particularmente los mamíferos, *por ejemplo* células COS 7, se pueden utilizar como expresión de las células huésped. Pequeños péptidos también pueden ser sintetizados en el laboratorio.

65

[0033] Con la disponibilidad de los polipéptidos en grandes cantidades, empleando un huésped de expresión, los polipéptidos se pueden aislar y purificar según formas convencionales. Un lisado se puede preparar del huésped de expresión y el lisado purificado usando HPLC, cromatografía de exclusión, electroforesis en gel, cromatografía de afinidad, u otra una técnica de purificación. El polipéptido purificado generalmente será por lo menos un 80% puro, preferentemente al menos un 90% puro, y puede ser hasta e incluyendo 100% puro. Puro se pretende que signifique libre de otras proteínas, así como restos celulares.

[0034] El polipéptido se utiliza para la producción de anticuerpos, donde los fragmentos cortos proporcionan anticuerpos específicos para el polipéptido particular, y fragmentos más grandes o toda la proteína permite la producción de anticuerpos sobre la superficie del polipéptido. Los anticuerpos se pueden elevar a péptidos aislados que corresponden a dominios particulares, o a la proteína nativa.

[0035] Los anticuerpos son preparados de acuerdo con las formas convencionales, donde se utiliza el polipéptido o proteína expresada como inmunógeno, por sí mismo o conjugado en portadores conocidos inmunogénicas, *por ejemplo* KLH, pre-S, HBsAg, otras proteínas virales o eucariotas, o similares. Se pueden utilizar varios adyuvantes, con una serie de inyecciones, como sea apropiado. Para anticuerpos monoclonales, después de una o más inyecciones de refuerzo, el bazo se aísla, los linfocitos se immortalizan mediante fusión celular, y luego se criban para una unión de anticuerpos de alta afinidad. Las células inmortalizadas, es decir hibridomas, que producen los anticuerpos deseados se pueden expandir a continuación. Para una descripción más detallada, consulte *Monoclonal Antibodies: A Laboratory Manual*, Harlow y Lane eds., Cold Spring Harbor Laboratories, Cold Spring Harbor, Nueva York, 1988. Si se desea, el ARNm que codifica las cadenas pesadas y ligeras puede aislarse y mutagenizarse mediante clonación en *E. coli*, y las cadenas pesadas y ligeras mezclarse para mejorar más la afinidad de los anticuerpos. Alternativas a la inmunización *in vivo* como procedimiento de aumento de anticuerpos incluyen la unión a librerías de "visualización" de fagos, usualmente en conjunción con la maduración de afinidad *in vitro*.

[0036] El anticuerpo se puede producir como una sola cadena, en lugar de la estructura normal de multímeros. Anticuerpos de cadena simple se describen en Jost et al. (1994) J.B.C. 269:26267-73, y otros. Las secuencias de ADN que codifican la región variable de la cadena pesada y la región variable de la cadena ligera se ligan a un espaciador que codifica por lo menos unos 4 aminoácidos de pequeños aminoácidos neutros, incluyendo glicina y/o serina. La proteína codificada mediante esta fusión permite el montaje de una región variable funcional que retiene la especificidad y la afinidad del anticuerpo original.

Utilización de marcadores de supresión en la identificación de micobacterias

[0037] La supresión Rv2660c proporcionada en la Tabla 1 es útil para la identificación de una micobacteria como (a) variantes de *M. tb.* (b) aislados de BCG (c) cepas de *M. bovis* o (d) llevando el bacteriófago micobacteriano identificado, dependiendo del marcador específico que se elija. Dicho cribado es particularmente útil para determinar si una infección particular o aislado es patógeno. El término micobacterias puede referirse a cualquier miembro de la familia Mycobacteriaceae, incluyendo *M. tuberculosis*, *M. avium complex*, *M. kansasii*, *M. scrofulaceum*, *M. bovis* y *M. leprae*.

[0038] Medios para detectar supresiones son conocidos en la técnica. Las supresiones se pueden identificar a través de la ausencia o la presencia de las secuencias de ARNm o ADN genómico, a través de análisis de las regiones de unión que flanquean la supresión, o la detección del producto génico, o, particularmente en relación a la prueba cutánea de la tuberculina, mediante la identificación de anticuerpos que reaccionan con el producto del gen codificado.

[0039] Aunque las supresiones se puede determinar fácilmente por la ausencia de hibridación, en muchos casos es deseable disponer de una señal positiva, con el fin de reducir al mínimo las lecturas negativas artificiales. En estos casos, las supresiones se pueden detectar mediante el diseño de unos cebadores que flanquean la unión formada por la supresión. Cuando la supresión está presente, se forma una nueva secuencia entre las regiones de flanco, que pueden detectarse mediante hibridación. Preferiblemente, este cebador será lo suficientemente corto que sólo se hibride con la unión, y dejará de formar híbridos estables con cualquiera de las partes separadas de la unión.

[0040] El diagnóstico se realiza mediante proteína, secuencia de ADN o ARN y/o análisis de hibridación de cualquier muestra conveniente, *por ejemplo* micobacterias cultivadas, material de biopsia, muestra de sangre, etc. El cribado también puede basarse en las características funcionales o antigénicas de la proteína. Los inmunoensayos diseñados para detectar las proteínas codificadas a partir de secuencias suprimidas pueden utilizarse en el cribado.

[0041] Una serie de procedimientos están disponibles para analizar de ácidos nucleicos para detectar la presencia de una secuencia específica. Si están disponibles grandes cantidades de ADN, el ADN genómico se utiliza directamente. Alternativamente, la región de interés es clonada en un vector adecuado y crece en una cantidad suficiente para su análisis. El ácido nucleico se puede amplificar mediante técnicas convencionales, tal como la reacción en cadena de la polimerasa (PCR), para proporcionar cantidades suficientes para el análisis. El uso de la reacción en cadena de la polimerasa se describe en Saiki, et al. (1985) *Science* 239:487, y una revisión de las técnicas actuales se pueden encontrar en Sambrook, et al. *Molecular Cloning: Laboratory Manual*, CSH Press 1989, pp. 14.2-14.33. La amplificación también se puede utilizar para determinar si un polimorfismo está presente, mediante el uso de un cebador que sea específico para el polimorfismo. Alternativamente, varios procedimientos son conocidos en la técnica que utilizan el

ligado de oligonucleótidos, para ejemplos ver Riley et al. (1990) N.A.R. 18:2887-2890, y Delahunty et al. (1996) Am. J. Hum. Genet. 58:1239-1246.

5 [0042] Una etiqueta detectable puede incluirse en una reacción de amplificación. Las etiquetas adecuadas incluyen fluorocromos, *por ejemplo* isotiocianato de fluoresceína (FITC), rodamina, Texas Red, ficceritrina, aloficocianina, 6-carboxifluoresceína (6-FAM), 2,7-dimetoxi-4',5'-dicloro-6-carboxifluoresceína (JOE), 6-carboxi-X-rodamina (ROX), 6-carboxi-2',4',7',4,7-hexaclorofluoresceína (HEX), 5-carboxifluoresceína (5-FAM) o N,N,N',N'-tetrametil-6-carboxi-rodamina (TAMRA), etiquetas radiactivas, *por ejemplo* ^{32}P , ^{35}S , ^3H : *etc.* La etiqueta puede ser un sistema de dos etapas, donde se
10 conjuga el ADN amplificado a biotina, haptenos, *etc.* que tienen un socio de unión de alta afinidad, *por ejemplo* avidina, anticuerpos específicos, *etc.*, donde se conjuga el socio de unión a una etiqueta detectable. La etiqueta puede conjugarse a uno o los dos cebadores. Alternativamente, el grupo de nucleótidos utilizados en la amplificación está etiquetado, para incorporar la etiqueta en el producto de amplificación.

15 [0043] La muestra de ácido nucleico, *por ejemplo* fragmento amplificado o clonado, es analizada mediante una de una serie de procedimientos conocidos en la técnica. El ácido nucleico puede ser secuenciado mediante dideoxi u otros procedimientos, y la secuencia de bases comparada con la secuencia suprimida. La hibridación con la secuencia variante también se puede usar para determinar su presencia, mediante Southern blot, dot blots, *etc.* El patrón de hibridación de una secuencia de control y la variante en una matriz de sondas de oligonucleótidos inmovilizadas en un soporte sólido, tal como se describe en los documentos US 5.445.934 o WO 95/35505, también puede utilizarse como
20 un medio para detectar la presencia de secuencias variables. Análisis de polimorfismo conformacional de cadena simple (SSCP), electroforesis en gel de gradiente desnaturante (DGGE), detección de división no coincidente, y análisis de heterodúplex en matrices de gel se utilizan para detectar cambios conformacionales creados por variación de la secuencia de ADN como alteraciones en la movilidad electroforética. Alternativamente, cuando un polimorfismo crea o destruye un sitio de reconocimiento para una endonucleasa de restricción (polimorfismo de longitud de los fragmentos de restricción, RFLP), la muestra se digiere con esa endonucleasa, y el tamaño de los productos se fracciona para
25 determinar si el fragmento fue digerido. El fraccionamiento se realiza mediante electroforesis en gel o capilaridad, en particular geles de acrilamida o agarosa.

30 [0044] El patrón de hibridación de una secuencia de control y variante en una matriz de sondas de oligonucleótidos inmovilizadas en un soporte sólido, tal como se describe en los documentos US 5.445.934 o WO 95/35505, puede utilizarse como un medio para detectar la presencia de secuencias suprimidas. En una realización de la invención, se proporciona una matriz de oligonucleótidos, donde las posiciones discretas de la matriz son complementarias por lo menos con una porción de ADN genómico de *M. tb.*, que usualmente comprende al menos una porción de los marcos de lectura abierta identificados. Esta matriz puede comprender una serie de oligonucleótidos, cada uno de los cuales
35 pueden hibridar específicamente en un ácido nucleico *por ejemplo* ARNm, ADNc, ADN genómico, *etc.*

[0045] Las supresiones también se pueden detectar mediante amplificación. En una realización de la invención, las secuencias se amplifican que incluyen una unión de supresión, es decir, cuando los cebadores de amplificación se hibridan con una secuencia de unión. En una muestra de ácido nucleico en la que la secuencia del marcador se suprime, se formará una unión, y el cebador se hibridará, permitiendo así la amplificación de una secuencia detectable. En una muestra de ácido nucleico en la que la secuencia del marcador está presente, el cebador no se hibridará, y no se realizará ninguna amplificación. Alternativamente, los cebadores de amplificación se pueden elegir de tal manera que la amplificación de la secuencia objetivo sólo se realizará si la secuencia del marcador está presente. Los productos de la amplificación se pueden separar por tamaños utilizando cualquier procedimiento conveniente, tal como se conoce en la técnica, incluyendo electroforesis en gel, cromatografía, electroforesis capilar, fraccionamiento de gradiente de densidad, *etc.*
40
45

[0046] Además de la detección de supresiones mediante la detección de secuencias de uniones, o la detección de las propias secuencias de marcado, se puede determinar la presencia o ausencia del producto de la proteína codificada. Las supresiones específicas en la tabla 1 se corresponden con marcos de lectura abierta del genoma de *M. tb.*, y por lo tanto codifican un polipéptido. Los polipéptidos se detectan mediante medios conocidos en la técnica, incluyendo la determinación de la presencia del polipéptido específico en una muestra a través de caracterización bioquímica, funcional o inmunológica. La detección de anticuerpos en el suero de pacientes que reaccionan con un polipéptido es de particular interés.
50
55

[0047] La inmunización con BCG habitualmente conduce a una respuesta positiva frente a los antígenos de la tuberculina en una prueba de la piel. En las personas que han sido inmunizadas, que incluye una proporción significativa de la población mundial, es por lo tanto difícil determinar si una prueba positiva es el resultado de una reacción inmune a la vacuna BCG, o de una infección de *M. tb.* en curso. La presente invención proporciona la Rv2660c que está presente en aislados de *M. tb.*, pero está ausente en BCG. Como procedimiento de cribado primario o secundario se puede analizar la inmunorreactividad del paciente con los polipéptidos codificados mediante dichos marcadores de supresión. El diagnóstico puede realizarse mediante una serie de procedimientos. Todos los diferentes procedimientos determinan la presencia de una respuesta inmune al polipéptido en un paciente, donde una respuesta positiva es indicativa de una infección de *M. tb.* La respuesta inmune se puede determinar mediante la determinación de la unión a anticuerpo, o mediante la presencia de una respuesta a una estimulación intradérmica con el polipéptido.
60
65

[0048] En un procedimiento, una dosis del polipéptido marcador de la supresión, formulado como un cóctel de proteínas o como especies de proteínas individuales, en un medio adecuado se inyecta por vía subcutánea en el paciente. La dosis será usualmente de al menos de 0,05 µg de proteína, y usualmente no mayor de 5 µg de proteína. Un control que comprende medio solamente, o una proteína no relacionada se inyectará cerca al mismo tiempo. El sitio de la inyección se examina después de un período de tiempo para la presencia de una roncha. La roncha en el sitio de la inyección del polipéptido se compara con que en el sitio de la inyección de control, usualmente midiendo el tamaño de la roncha. Las lecturas de prueba de la piel se pueden evaluar mediante una variedad de sistemas de clasificación objetivos. Un resultado positivo de la presencia de una condición alérgica mostrará un diámetro mayor en el sitio de la inyección del polipéptido en comparación con el control, usualmente al menos un 50% de aumento en el tamaño, más usualmente un aumento de por lo menos un 100% en el tamaño.

[0049] Un procedimiento alternativo para el diagnóstico depende de la detección *in vitro* de la unión entre anticuerpos en una muestra del paciente y los polipéptidos objetivo, ya sea como un cóctel o como especies de proteínas individuales, donde la presencia de unión específica es indicativa de una infección. La medición de la concentración de anticuerpos específicos de polipéptido en una muestra o fracción del mismo puede lograrse mediante una variedad de ensayos específicos. En general, el ensayo medirá la reactividad entre una muestra del paciente, usualmente derivados de sangre, generalmente en forma de plasma o suero. La muestra del paciente se puede utilizar directamente o diluida como se apropiado, usualmente 1:10 aproximadamente y usualmente no mayor de 1:10.000 aproximadamente. Los inmunoensayos se pueden realizar en cualquier tampón fisiológico, *por ejemplo* PBS, solución salina normal, HBSS, DPBS, *etc.*

[0050] En una realización preferida, se utiliza un ensayo de tipo sándwich convencional. Un ensayo sándwich se realiza en primer lugar fijando el polipéptido a una superficie o soporte insoluble. El polipéptido puede unirse a la superficie mediante cualesquiera medios convenientes, dependiendo de la naturaleza de la superficie, ya sea directamente o por medio de anticuerpos específicos. La forma particular de unión no es crucial, siempre que sea compatible con los reactivos y los procedimientos generales de la invención. Se pueden unir a las placas de forma covalente o no covalente, preferiblemente de manera no covalente. Las muestras, fracciones o partes alícuotas del mismo se añaden a continuación a los soportes por separado de manera que se puedan ensayar (por ejemplo, pocillos separados por una placa de microtitulación) que contienen polipéptido de unión de soporte. Preferiblemente, una serie de estándares, que contienen concentraciones conocidas de anticuerpos, se analizan en paralelo con las muestras o partes alícuotas de las mismas para servir de controles.

[0051] Los receptores inmunes específicos se pueden etiquetar para facilitar la cuantificación directa o indirecta de la unión. Ejemplos de etiquetas que permiten la medición directa de una segunda unión del receptor incluyen radioetiquetas, tal como ^3H o ^{125}I , agentes fluorescentes, tintes, cuentas, quimioluminescentes, partículas coloidales, y similares. Ejemplos de etiquetas que permiten la medición indirecta de la unión incluyen enzimas donde el sustrato puede proporcionar un producto coloreado o fluorescente en una realización preferida, los segundos receptores son anticuerpos marcados con una enzima unida de manera covalente capaces de proporcionar una señal de producto detectable después de la adición del sustrato adecuado. Ejemplos de enzimas adecuadas para su uso en conjugados incluyen peroxidasa de rábano, fosfatasa alcalina, malato deshidrogenasa y similares. Cuando no están disponibles comercialmente, estos conjugados de anticuerpos-enzimas se producen fácilmente mediante técnicas conocidas por los expertos en la materia.

[0052] En algunos casos, se utilizará un ensayo competitivo. Además de la muestra del paciente, se agrega un competidor del anticuerpo a la mezcla de reacción. El competidor y el anticuerpo compiten para unirse al polipéptido. Usualmente, la molécula competidora se marcará y detectará tal como se describió anteriormente, cuando la cantidad de unión del competidor sea proporcional a la cantidad de inmune presente. La concentración de la molécula del competidor será de aproximadamente 10 veces la máxima concentración inmune anticipada a aproximadamente la misma concentración para hacer el rango de detección más sensible y lineal.

[0053] Alternativamente, los anticuerpos pueden ser utilizados para la determinación directa de la presencia del polipéptido marcador de supresión. Anticuerpos específicos para los marcadores de supresión del sujeto tal como se han descrito pueden utilizarse en inmunoensayos de cribado. Las muestras, tal como se usan aquí, incluyen cultivos microbianos, fluidos biológicos, tales como lavado traqueal, sangre, *etc.* También se incluyen en el término los derivados y las fracciones de estos líquidos. El diagnóstico se puede realizar mediante una serie de procedimientos. Los procedimientos diferentes determinan todos la ausencia o la presencia de polipéptidos codificados mediante los marcadores de supresión del sujeto. Por ejemplo, la detección puede utilizar la tinción de células micobacterianas o secciones histológicas, realizado de acuerdo con procedimientos convencionales. Los anticuerpos de interés se agregan a la muestra celular, y se incuban durante un período de tiempo suficiente para permitir la unión al epítipo, generalmente por lo menos 10 minutos aproximadamente. El anticuerpo puede marcarse con radioisótopos, enzimas, agentes fluorescentes, quimioluminescentes, u otras etiquetas para la detección directa. Alternativamente, un anticuerpo de segunda etapa o reactivo se utiliza para amplificar la señal. Estos reactivos son bien conocidos en la técnica. Por ejemplo, el anticuerpo primario puede ser conjugado a biotina, con avidina conjugada con peroxidasa de rábano picante añadido como reactivo de segunda etapa. La detección de final utiliza un sustrato que se somete a un cambio de color en presencia de la peroxidasa. La ausencia o presencia de la unión del anticuerpo se puede determinar mediante varios procedimientos, incluyendo microscopía, radiografía, recuento de centelleo, *etc.*

[0054] Un procedimiento alternativo para el diagnóstico depende de la detección *in vitro* de la unión entre anticuerpos y los polipéptidos del sujeto en solución, por ejemplo, un lisado celular. La medición de la concentración de la unión en una muestra o fracción de la misma puede lograrse mediante una variedad de ensayos específicos. Se puede utilizar un ensayo de tipo sándwich convencional. Por ejemplo, un ensayo sándwich puede primero fijar anticuerpos específicos a una superficie o soporte insoluble. La forma particular de unión no es crucial, siempre que sea compatible con los reactivos y los procedimientos generales de la invención. Se pueden unir a las placas de forma covalente o no covalente, preferiblemente no covalente. Los soportes insolubles pueden ser cualesquiera composiciones a las que se pueden unir polipéptidos, que se separan fácilmente del material soluble y, que por otro lado es compatible con el procedimiento global. La superficie de dichos soportes puede ser sólida o porosa y de cualquier forma conveniente. Ejemplos de soportes insolubles adecuados a los que está unido el receptor incluyen cuentas, *por ejemplo* cuentas magnéticas, membranas y placas de microtitulación. Típicamente son de vidrio, plástico (*por ejemplo* poliestireno), polisacáridos, nylon o nitrocelulosa. Las placas de microtitulación son especialmente convenientes debido a que un gran número de ensayos pueden llevarse a cabo simultáneamente, utilizando pequeñas cantidades de reactivos y muestras.

[0055] Las muestras se añaden a continuación a los soportes que se pueden ensayar por separado (por ejemplo, pozos separados de una placa de microtitulación) que contienen anticuerpos. Preferiblemente, una serie de estándares, que contienen concentraciones conocidas de los polipéptidos se ensayan en paralelo con las muestras o partes alícuotas de las mismas para servir como controles. Preferiblemente, cada muestra y estándar se añadirá a los múltiples pocillos, de modo que se pueden obtener para cada uno los valores promedio. El tiempo de incubación debe ser suficiente para la unión, en general, de 0,1 a 3 horas aproximadamente es suficiente. Después de la incubación, el soporte insoluble se suele lavar de componentes de no unión. Generalmente, un medio detergente no iónico diluido a un pH adecuado, generalmente 7-8, se utiliza como medio de lavado. Se pueden utilizar de uno a seis lavados, con un volumen suficiente para lavar completamente las proteínas no específicamente unidas presentes en la muestra.

[0056] Después del lavado, se aplica una solución que contiene un segundo anticuerpo. El anticuerpo se unirá con la especificidad suficiente de modo que pueda distinguirse de otros componentes presentes. Los segundos anticuerpos se pueden etiquetar para facilitar la cuantificación directa o indirecta de la unión. Ejemplos de etiquetas que permiten la medición directa de la unión del segundo receptor incluyen radioetiquetas, tal como ^3H o ^{125}I , agentes fluorescentes, tintes, cuentas, quimioluminiscentes, partículas coloidales, etc. Ejemplos de etiquetas que permiten la medición indirecta de aproximaciones incluyen enzimas donde el sustrato puede proporcionar un producto de color o fluorescente. En una realización preferida, los anticuerpos son marcados con una enzima covalentemente unida capaz de proporcionar una señal de producto detectable después de la adición del sustrato adecuado. Ejemplos de enzimas adecuadas para su uso en conjugados incluyen peroxidasa de rábano, fosfatasa alcalina, malato deshidrogenasa y similares. Si no están comercialmente disponibles, estos conjugados de anticuerpo-enzima se producen fácilmente por técnicas conocidas por los expertos en la materia. El tiempo de incubación debe ser suficiente para el ligando marcado para unirse a las moléculas disponibles. Generalmente, de 0,1 a 3 horas aproximadamente es suficiente, usualmente 1 hora es suficiente.

[0057] Después de la segunda etapa de unión, el soporte insoluble se lava de nuevo libre de material no específicamente unido. La señal producida mediante el conjugado unido se detecta mediante medios convencionales. Cuando se utiliza un conjugado de enzima, se proporciona un sustrato de enzima apropiado de manera que se forma un producto detectable.

[0058] Otros inmunoensayos son conocidos en la técnica y puede encontrar uso como diagnóstico. Las placas Ouchterlony ofrecen una determinación simple de la unión del anticuerpo. Western blots se pueden realizar en geles de proteínas o puntos de proteínas en filtros, utilizando un sistema de detección específico para el polipéptido, convenientemente usando un procedimiento de etiquetado tal como se describe para el ensayo de sándwich.

50 Mycobacterium Recombinante

[0059] *Mycobacterium*, en particular las del complejo *M. tuberculosis*, están diseñadas genéticamente para contener supresiones o inserciones específicas correspondientes a los marcadores genéticos identificados. En particular, las cepas atenuadas de BCG se modifican para introducir genes suprimidos que codifican las secuencias importantes en el establecimiento de una inmunidad efectiva. Alternativamente, *M. bovis* o *M. tuberculosis* se modifican mediante recombinación homóloga para crear supresiones específicas en secuencias que determinan la virulencia, es decir, las bacterias se atenúan mediante técnicas recombinantes.

[0060] Para introducir de manera estable secuencias en BCG, el marco de lectura abierto de *M. tb* correspondiente a Rv2660c se inserta en un vector que se mantiene en cepas de *M. bovis*. Preferentemente, se incluyen secuencias flanqueantes 5' y 3' nativas, para proporcionar una regulación adecuada de la transcripción y traducción. Sin embargo, en circunstancias especiales, pueden incluirse promotores exógenos y otras regiones reguladoras. Los vectores y los procedimientos de transfección para BCG se conocen en la técnica. Por ejemplo, la patente US 5.776.465 describe la introducción de genes exógenos en BCG.

- 5 [0061] En una realización de la invención, la región completa suprimida es sustituida en BCG. Las uniones de la supresión se determinan en relación con un tipo salvaje de secuencia de *M. tb.* o *M. bovis*, por ejemplo, tal como se establece en la sección experimental. La región suprimida se clona mediante cualquier procedimiento conveniente, tal como se conoce en la técnica, *por ejemplo* amplificación por PCR de la región, digestión con endonucleasas de restricción, síntesis química, *etc.* Preferiblemente, la región clonada también comprenderá secuencias de flaqueo de una longitud suficiente para inducir la recombinación homóloga, usualmente por lo menos 25 nt aproximadamente, más usualmente por lo menos 100 nt aproximadamente, o mayor. Vectores y procedimientos adecuados son conocidos en la técnica, por ejemplo, véase Norman et al. (1995) Mol. Microbiol. 16:755-760.
- 10 [0062] En una realización alternativa se introduce Rv2660c en una cepa de *M.tuberculosis* o *M.bovis*. Preferiblemente, dicha cepa se reduce en virulencia, por ejemplo H37Ra, *etc.* En la técnica se conocen procedimientos de recombinación homóloga para realizar supresiones en micobacterias, por ejemplo, véase Norman *et al.*, *supra.*; Ganjam et al. (1991) P.N.A.S. 88:5433-5437, y Aldovini et al. (1993) J. Bacteriol. 175:7282-7289. Las supresiones pueden comprender un marco de lectura abierto identificado en la Tabla 1, o pueden extenderse a la supresión completa, es decir, se extiende a las regiones flanqueantes, y pueden incluir múltiples marcos de lectura abierta.
- 15 [0063] La capacidad de la micobacteria genéticamente alterada de causar enfermedades se puede probar en uno o más modelos experimentales. Por ejemplo, *M. tb.* se sabe que infecta a una variedad de animales, y células en cultivo. En un ensayo, se infectan macrófagos de mamíferos, preferentemente macrófagos humanos. En una comparación de cepas virulentas, no virulentas y atenuadas del complejo *M. tuberculosis*, monocitos alveolares o de sangre periférica se infectan en una proporción 1:1 (Silver et al. (1998) Infect Immun 66 (3):1190-1199; Paul et al (1996) J Infect Dis 174(1) 105-112). Los porcentajes de células infectadas mediante las cepas y el número inicial de microorganismos intracelulares son equivalentes, al igual que los niveles de la viabilidad de monocitos hasta 7 días después de la infección. Sin embargo, el crecimiento intracelular refleja la virulencia, en un periodo de una o más semanas. El crecimiento de micobacterias se puede evaluar mediante tinción de ácidos rápidos, microscopía electrónica, y ensayo de unidades formadoras de colonias (ufc). La producción de monocitos de factor alfa de necrosis tumoral también puede monitorizarse como un marcador de la virulencia.
- 20 [0064] Otros ensayos de la virulencia utilizan modelos animales. Las bacterias del complejo *M. tuberculosis* son capaces de infectar una amplia variedad de huéspedes animales. Un modelo de especial interés es la tuberculosis cavitaria producida en conejos por bacilos tuberculosos virulentos en aerosol (Converse et al. (1996) Infect Immun 64 (11):4776-4787). En caseosa licuada, el bacilo de la tuberculosis crece extracelularmente desde el primer momento del inicio de la enfermedad y puede llegar a un número tan grande que se pueden desarrollar mutantes con resistencia a los antimicrobianos. Desde una cavidad, los bacilos entran en el árbol bronquial y se extienden a otras partes del pulmón y también a otras personas. De los animales de laboratorio de uso común, el conejo es el único en que se puede producir fácilmente la tuberculosis cavitaria.
- 30 [0065] Las vacunas se pueden formular de acuerdo a procedimientos conocidos en la técnica. Las vacunas de la bacteria modificada se administran a un huésped que puede estar expuesto a la tuberculosis virulenta. En muchos países donde la tuberculosis es endémica, la vacunación puede administrar al nacer, con vacunas adicionales si es necesario. Los compuestos de la presente invención se administran en una dosis que proporciona eficacia y minimiza cualquier efecto secundario. Se contempla que la composición se obtendrá y se utilizará bajo la guía de un médico.
- 35 [0066] Las cepas de vacuna convencionales de BCG se pueden formular en una vacuna combinada con polipéptidos identificados en la presente invención y producidas tal como se describió anteriormente, para mejorar la eficacia de la vacuna.
- [0067] Varios procedimientos para la administración pueden ser empleados. La formulación puede inyectarse por vía intramuscular, por vía intravascular, por vía subcutánea, *etc.* La dosis será convencional. La bacteria se puede formular en composiciones farmacéuticas mediante combinación con los portadores o diluyentes farmacéuticamente aceptables apropiados, y se pueden formular en preparaciones en formas semisólidas o líquidas, tales como soluciones, inyecciones, *etc.* Los siguientes procedimientos y excipientes son meramente de ejemplo y no son de ninguna manera restrictivos.
- 50 [0068] La bacteria modificada se puede formular en preparaciones para inyecciones mediante disolución por suspensión o emulsión de los mismos en un solvente acuoso o no acuoso, como otros aceites vegetales o similares, glicéridos de ácidos alifáticos sintéticos, ésteres de ácidos alifáticos superiores o propileno glicol, y si se desea, con aditivos convencionales tal como solubilizantes, agentes isotónicos, agentes en suspensión, agentes emulsionantes, estabilizantes y conservantes. Las formas de dosificación unitarias para la inyección o la administración intravenosa pueden comprender la bacteria de la presente invención en una composición tal como una solución en agua estéril, salino normal u otro portador farmacéuticamente aceptable.
- 60 [0069] El término "forma de dosificación unitaria" tal como se usa aquí, se refiere a unidades físicamente discretas adecuadas como dosis unitarias para humanos y animales, cada unidad contiene una cantidad predeterminada de la vacuna, calculada en una cantidad suficiente para producir el efecto deseado en asociación con un diluyente, portador o vehículo farmacéuticamente aceptable. Las especificaciones para las formas de dosificación unitarias de la presente
- 65

invención dependerán de la bacteria utilizada en concreto y el efecto que se desee alcanzar, y la farmacodinamia asociada a cada complejo en el huésped.

5 [0070] Los excipientes farmacéuticamente aceptables, tales como vehículos, adyuvantes, portadores o diluyentes, están disponibles al público. Por otra parte, las sustancias auxiliares farmacéuticamente aceptables, tales como el pH de ajuste y los agentes de tampón, ajustando los agentes de tonicidad, estabilizadores, agentes humectantes y similares, están fácilmente disponibles para el público.

10 EXPERIMENTAL

Procedimientos:

15 [0071] Los procedimientos técnicos utilizados comienzan con la extracción de ADN genómico completo de bacterias que crecen en cultivo.

Día 1

20 [0072] Se inocula medio de cultivo de elección (LJ/7H9) y se incuba a 35°C hasta que un crecimiento abundante de Dispense 500 µl 1x TE en cada tubo. (Si el ADN se encuentra en medio líquido, no es necesario TE). Se transfiere el contenido de una asa de siembra (sedimento) de las células en un tubo de microcentrifuga que contiene 500µl de 1*TE. Si se toma ADN del medio líquido, se deja que las células se acumulen en la parte inferior del frasco. Se pipetea las células (1 ml aproximadamente) en el tubo. Se calientan 20 min a 80°C para matar las células, se centrifugan, se resuspenden en 500µl de 1*TE. Se añaden 50µl de 10 mg/ml de lisozima, se agitan en vórtice, y se incuban durante una noche a 37°C.

25 Día 2

30 [0073] Se añaden 70µl de SDS al 10% y 10 µl de proteinasa K y se agitan e incuban durante 20 min. a 65°C. Se añaden 100µl de 5 M NaCl. Se agregan 100µl de CTAB/NaCl, previamente calentado a 65°C. Se agitan hasta un líquido de contenido blanco ("leche"). Se incuban durante 10 min a 65°C. Fuera de la campana, se preparan nuevos tubos de microcentrifuga etiquetados con el # de cultivo en la parte superior, y el # de cultivo, el # de tubo, la fecha en el lateral. Se añaden 550 microlitros con isopropanol a cada uno y se tapan. De vuelta a la campana, se incrementa en 750 l de cloroformo/alcohol isoamílico y se agita durante 10 seg. Se centrifuga a temperatura ambiente durante 5 min. a 12.000 g. El sobrenadante acuoso de transferencia en cantidades de 180µl asciende a nuevo tubo usando una pipeta, teniendo cuidado de dejar atrás los sólidos y los líquidos no acuosos. Se ponen durante 30 minutos a -20°C. Se agitan durante 15 min a temperatura ambiente en una microcentrifuga a 12.000 g. Se desecha el sobrenadante, y se dejan alrededor de 20 µl por encima de precipitado. Se añade 1 ml de etanol frío al 70% y se gira el tubo un par de veces de arriba a abajo. Se agita durante 5 min a temperatura ambiente en una microcentrifuga. Se desecha el sobrenadante, se dejan alrededor de 20µl por encima del precipitado. Se agita durante 1 min en una microcentrifuga y se desecha con cuidado los últimos 20 µl de sobrenadante, justo por encima del precipitado con una pipeta (P-20). Asegúrese de que todas las trazas de etanol se quitan. Las cuentas se dejan secar a temperatura ambiente durante 10 minutos o velocidad de vacío de 2-3 min. (Se colocan los tubos abiertos en la velocidad de vacío, se cierra la tapa, se arranca el rotor, se activa el vacío. Después de 3 min. se oprime el botón rojo, se desactiva el vacío, se apaga el rotor. Se comprueba si están secas las cuentas mediante el tubo para ver si se eliminan las cuentas de un lado del tubo). Se recoge el precipitado en 20-50 µl de ddH2O. Las cuentas pequeñas obtienen 20, de tamaño regular obtienen 30 y las muy grandes obtienen 50. El ADN se puede almacenar a 4°C para su posterior utilización.

45 [0074] *Matriz de ADN:* Se realizó mediante la colocación de fragmentos de ADN en portaobjetos de vidrio que fueron tratados previamente con poli-L-lisina. La colocación sobre la matriz se realizó mediante un colocador robótico. El ADN se reticuló con el vidrio mediante irradiación ultravioleta, y los grupos libres de poli-L-lisina fueron bloqueados mediante tratamiento con 0,05% de anhídrido succínico, 50% de 1-metil-2-pirrolidinona y 50% de tampón de borato.

55 [0075] La mayoría de los puntos de la matriz de productos derivados de PCR, producidos mediante la selección de más de 9000 pares de cebadores diseñados para amplificar los marcos de lectura abierta predichos de las secuencias de la cepa H37Rv (<ftp.sanger.ac.uk/pub/TB.seq>). Algunas estándares internos y puntos negativos de control incluyen vectores de plásmido y ADN no-*M.tb.* también estuvieron presentes en la matriz.

60 [0076] Por lo tanto, con la preparación de una matriz que contenía la totalidad del genoma de *Mycobacterium tuberculosis*, que comparó BCG-Connaught con *Mycobacterium tuberculosis*, utilizando la matriz para la hibridación competitiva. El protocolo siguiente:

65 [0077] *Protocolo de etiquetado de ADN.* Añadir 4 mg de ADN en 20 µl de H2O, 2 ml de dN10N6 y 36 µl de H2O. 2 ml de ADN pico por cada muestra de ADN, para el total de 60µl. Hervir 3 minutos para desnaturalizar el ADN, luego presión fría en baño de agua helada. Añadir 1 µl de dNTP (5 mM ACG), 10 µl de tampón 10,4 µl de Klenow, 22 µl de H2O en cada tubo. Añadir 3 µl de Cy3 o Cy5 dUTP, para un total de 100 µl. Incubar 3 horas a 37C. Añadir 11µl de 3M NAAC, 250 µl de EtOH 100% para precipitar, almacenar O/N a -20C. Centrifugar las muestras genómicas 30 minutos a 13K

para que sedimente el precipitado. Eliminar el sobrenadante, añadir 70% EtOH, agitar 15 minutos, descarga sup. y velocidad de vacío para secar. Esto proporciona ADN para dos experimentos.

5 [0078] *Protocolo de hibridación de ADN en micromatrices.* Resuspender el ADN marcado en 11 µl de dH₂O (por 2 matrices). Retirar 1 µl de ADN en un gel de agarosa al 1,5% del documento de muestra a hibridar. De los restantes 10 µl de solución, la mitad se utilizará para esta hib, y la otra mitad quedará para más adelante. Tome 5µl de Cy3 en solución y añadir la misma cantidad de solución Cy5, para el volumen total de 10 µl de ADN marcado mezclado. Añadir 1 µl de ARNt, 2,75 µl de 20x SSC, 0,4 µl de SDS, para un volumen total de 14,1 µl. Coloque en el portaobjetos en el lugar de la matriz, cubrir con el cubreobjetos de 22 mm, poner el portaobjetos de vidrio por encima y apretar sobre dispositivos de caucho, a continuación hibridar 4 horas a 65C. Después de 4 horas, retirar los portaobjetos de matriz de los dispositivos, dejar el cubreobjetos e introducirlo en la bandeja portaobjetos en tampón de lavado que consiste en 1x SSC con SDS 0,05% durante 2 minutos. El cubreobjetos debe caerse en el baño. Después de 2 minutos en tampón de lavado, sumergir una vez en un baño con 0,06x SSC, a continuación, volver a enjuagar con 0,06x SSC en un baño separado. Secar los portaobjetos en centrifuga de 600 rpm aproximadamente. Ahora están listos para la exploración.

10 [0079] *Exploración de fluorescencia y adquisición de datos.* La exploración de fluorescencia se fijó para 20 micras/píxel y dos lecturas fueron tomadas por píxel. Los datos para el canal 1 se crearon para recoger la fluorescencia de Cy3 con excitación a 520 nm y emisión a 550-600 nm. Las señales del canal 2 conseguidas se excitaron a 647 nm y se emitieron a 660-705 nm, apropiado para Cy5. Ningún filtro de densidad neutra se aplicó a la señal de ningún canal, y la ganancia del tubo fotomultiplicador se establece en 5. Se realizaron a continuación ajustes finos de la ganancia del fotomultiplicador para que las señales recogidas de los dos puntos que contiene el ADN genómico fueran equivalentes.

15 [0080] Para analizar la señal de cada punto en la matriz, una cuadrícula de cajas de 14x14 se aplicó a los datos obtenidos de la matriz, de tal manera que las señales dentro de cada caja se integraron y un valor se asignó al punto correspondiente. Un valor de fondo se obtuvo para cada punto mediante la integración de las señales medidas 2 píxeles fuera del perímetro de la casilla correspondiente. La señal y los valores básicos para cada punto se importaron a un programa de hoja de cálculo para su posterior análisis. Los valores de fondo se restaron de las señales y un factor de 1,025 se aplicó a cada valor en el canal 2 para normalizar los datos respecto a las señales de los puntos de ADN genómico.

20 [0081] Debido a que las dos muestras se etiquetan con diferentes tintes fluorescentes, es posible determinar que un punto de ADN en la matriz se ha hibridado con *Mycobacterium tuberculosis* (tinte verde) y no con BCG (tinte rojo), lo que demuestra una probable supresión del genoma BCG.

25 [0082] Sin embargo, como la matriz contiene ahora puntos que representan 4.000 puntos, se pueden esperar hasta 100 puntos con la hibridación de dos desviaciones estándar por encima o por debajo de la media. En consecuencia, hemos previsto un protocolo de cribado, en donde buscamos la hibridación no coincidente en dos genes consecutivos en el genoma. Por lo tanto, estamos esencialmente buscando sólo supresiones de genes múltiples en este momento.

30 [0083] Para confirmar que un gen o grupo de genes se suprime, realizamos hibridación Southern, empleando una sonda separada del ADN en la matriz. Las digestiones de ADN de *Mycobacterium* diferentes se ejecutan en un gel de agarosa, y se transfieren a las membranas. Las membranas pueden ser utilizadas repetidamente cribando para diferentes secuencias de ADN. A los efectos de este proyecto, se incluye el ADN de la cepa de referencia de *Mycobacterium tuberculosis* (H37Rv), a partir de otras cepas de laboratorio, tales como H37Ra, la cepa O, a partir de los aislados clínicos, a partir de la cepa de referencia de *Mycobacterium bovis*, y de diferentes cepas de *Mycobacterium bovis* BCG.

35 [0084] Una vez que se confirma una supresión mediante hibridación meridional, entonces se propuso caracterizar la localización genómica exacta. Esto se hace mediante la reacción en cadena de la polimerasa, con cebadores diseñados para estar cerca de los bordes de la supresión, véase Talbot (1997) J Clin Micro. 35: 566-9.

40 [0085] Se han elegido cebadores para amplificar a través de la región suprimida. Sólo en ausencia de esta región uno obtiene un amplicón. Los productos PCR fueron examinados por electroforesis (agarosa al 1,5%) y tinción con bromuro de etidio.

45 [0086] Una vez que se obtiene un amplicón corto, esta amplificación se secuencia a continuación. Se realiza una búsqueda en la base de datos del genoma para determinar si la secuencia es exactamente idéntica a una parte del genoma de *Mycobacterium tuberculosis*, y que la siguiente parte de la amplificación es exactamente idéntica a la otra parte del genoma de *Mycobacterium tuberculosis*. Esto permite una identificación precisa del sitio de supresión.

50 [0087] A continuación un ejemplo del tipo de informe obtenido:
rd6 unido a PCR, búsqueda blast de la secuencia:

emb|Z79701|MTCY277 Cósmido Mycobacterium tuberculosis Y277

Longitud = 38.908

Más Cadena HSPs:

Puntuación=643 (177,7 bits), Esperado=1,6e-54, Sum P(2) = 1,6e-54

Identicidades = 129/131 (98%), Positivas = 129/131 (98%), Cadena = Más /

Más

Consulta: 12 **ANTAGTAATGTGCGAGCTGAGCGATGTCGCCGCTCCCAAAAATTACCAATGGTTNGGTCA**
71

| ||| |

Sujeto:24784 **AGTAGTAATGTGCGAGCTGAGCGATGTCGCCGCTCCCAAAAATTACCAATGGTTTGGTCA**

Consulta: 72 **TGACGCCTTCCTAACCGAATTGTGAATTCATACAAGCCGTAGTCGTGCAGAAGCGCAAC**

| | | | |

Sujeto:24844 **TGACGCCTTCCTAACCGAATTGTGAATTCATACAAGCCGTAGTCGTGCAGAAGCGCAAC**

Consulta: 132 **ACTCTTGGAGT** 142

| | | | |

Sujeto: 24904 **ACTCTTGGAGT** 24914

Punt. = 224 (61.9 bits), Esper. = 1.6e-54, Sum P(2) = 1.6e-54

Identicidades= 46/49 (93%), Positivas= 46/49 (93%), Cadena = Más/Más

Consulta: 141 **GTGGCCTACAACGGNGCTCTCCGNGGCGGGCGGTACCGGATATCTTAG** 189

| | | | |

Sujeto: 37645 **GCGGCCTACAACGGCGCTCTCCGCGGCGGGCGGTACCGGATATCTTAG** 37693

Este proceso se repite con cada supresión propuesta, comenzando con las tres supresiones descritas previamente para servir como controles. 16 supresiones han sido identificadas mediante estos procedimientos, y se enumeran en la Tabla 5 1.

[0088] Debe entenderse que esta invención no se limita a la metodología particular, protocolos, fórmulas y reactivos descritos, ya que pueden, por supuesto, variar. También debe entenderse que la terminología utilizada en este documento es con el propósito de describir realizaciones particulares solamente, y no tiene por objeto limitar el ámbito de aplicación de la presente invención, que se limitará sólo por las reivindicaciones adjuntas. 10

[0089] Debe tenerse en cuenta que como se usa aquí y en las reivindicaciones adjuntas, la forma singular "a", "y", y "el/la" incluyen referentes plurales a menos que el contexto claramente indiquen otra cosa. Así, por ejemplo, la referencia a "un complejo" incluye una pluralidad de estos complejos y la referencia a "la formulación" incluye una referencia a una o varias formulaciones y equivalentes de la misma conocidas por los expertos en la materia, y así sucesivamente. 15

[0090] A menos que se defina de otro modo, todos los términos técnicos y científicos utilizados en este documento tienen el mismo significado que se entiende comúnmente por los expertos en la materia a la que pertenece esta invención. Aunque los procedimientos, aparatos y materiales, similares o equivalentes a los aquí descritos pueden ser utilizados en la práctica o en el ensayo de la invención, los procedimientos, dispositivos y materiales preferidos se describen a continuación. 20

[0091] Las publicaciones indicadas anteriormente y en todo el texto se proporcionan únicamente para su divulgación antes de la fecha de presentación de la presente solicitud. Nada aquí se interpretará como una admisión de que los inventores no tienen derechos anteriores a dicha divulgación, en virtud de la invención previa. 25

REIVINDICACIONES

1. Composición inmunógena, que comprende un polipéptido sustancialmente puro codificado por una secuencia de nucleótidos que comprende el marco de lectura abierto Rv2660c de *M. tuberculosis*:
 5 o un polipéptido codificado por un fragmento de nucleótidos de por lo menos 25 nucleótidos contiguos del marco de lectura abierto Rv2660c de *M. tuberculosis*;
 o dicho polipéptido fusionado a otras proteínas o partes de las mismas;
 y un excipiente farmacéuticamente aceptable.
- 1 0 2. Composición inmunógena según la reivindicación 1, que comprende además un adyuvante.
3. Composición inmunógena según la reivindicación 1, en la que dicho polipéptido está formulado conjuntamente con una micobacteria del complejo de *M. tuberculosis*.
- 1 5 4. Composición inmunógena según la reivindicación 3, en la que dicha micobacteria del complejo de *M. tuberculosis* es bacilo Calmette-Guerin o *M. bovis*.
5. Composición inmunógena según la reivindicación 1, para utilizar en un método de inmunización de un individuo a *M. tuberculosis*.
- 2 0 6. Micobacteria alterada genéticamente del complejo de *M. tuberculosis*, que comprende una secuencia de ácidos nucleicos exógena que comprende el marco de lectura abierto Rv2660c de *M. tuberculosis* o por lo menos 25 nucleótidos contiguos del marco de lectura abierto Rv2660c de *M. tuberculosis*.
- 2 5 7. Micobacteria alterada genéticamente según la reivindicación 6, en la que dicho ácido nucleico exógeno codifica un polipéptido que se fusiona a otro péptido o proteína.
8. Micobacteria alterada genéticamente según la reivindicación 6, en la que dicha micobacteria es BCG.
- 3 0 9. Micobacteria según la reivindicación 6, y un portador fisiológicamente aceptable para la inyección.
10. Método para determinar si un paciente se ha expuesto a un miembro del complejo de la tuberculosis, comprendiendo el método:
 3 5 poner en contacto una muestra del paciente con un polipéptido codificado por por lo menos 25 nucleótidos del marco de lectura abierto Rv2660c de *M. tuberculosis*, cuyo polipéptido no presenta reactividad cruzada para la exposición a BCG;
 y determinar la presencia de una reacción inmune a dicho polipéptido, donde una respuesta positiva es indicativa de exposición al complejo de tuberculosis.
- 4 0 11. Método según la reivindicación 10, en el que el miembro del complejo de tuberculosis es *Mycobacterium tuberculosis* o *Mycobacterium bovis*.
12. Método según la reivindicación 10, en el que dicha muestra comprende una muestra de sangre o un derivado de la misma.
- 4 5 13. Método según la reivindicación 10, en el que dicha etapa de determinación comprende: detectar la unión de un anticuerpo a dicho polipéptido, siendo dicha unión una indicación de que dicho sujeto está infectado por *Mycobacterium tuberculosis* o ha enfermado con *Mycobacterium tuberculosis*.
- 5 0 14. Polipéptido codificado por el marco de lectura abierto Rv2660c de *M. tuberculosis* para su uso en un método de diagnóstico *in vivo*.
15. Polipéptido tal como se define en la reivindicación 14 para su uso en el diagnóstico de la tuberculosis causada por un miembro del complejo de tuberculosis.
- 5 5