

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 809 167**

51 Int. Cl.:

C07K 14/285 (2006.01)

A61K 39/012 (2006.01)

C07K 16/12 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **17.01.2007 PCT/SE2007/000034**

87 Fecha y número de publicación internacional: **26.07.2007 WO07084053**

96 Fecha de presentación y número de la solicitud europea: **17.01.2007 E 07701118 (7)**

97 Fecha y número de publicación de la concesión europea: **24.06.2020 EP 1973933**

54 Título: **Una novedosa proteína de Haemophilus influenzae expuesta en la superficie (proteína E; pE)**

30 Prioridad:

17.01.2006 US 758987 P

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

03.03.2021

73 Titular/es:

**FORSGREN, ARNE (100.0%)
Sothönsvägen 4 B
230 11 Falsterbo, SE**

72 Inventor/es:

RIESBECK, KRISTIAN

74 Agente/Representante:

ISERN JARA, Jorge

ES 2 809 167 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

Una novedosa proteína de *Haemophilus influenzae* expuesta en la superficie (proteína E; pE)

5 Campo de la invención

La presente invención se refiere a la proteína E expuesta en la superficie, un factor de virulencia, que existe en todas las *Haemophilus influenzae* encapsuladas y no tipificables.

10 Antecedentes de la invención

Tanto la *Haemophilus influenzae* tipo b (Hib) como la *H. influenzae* no tipificable (NTHi) causan una variedad de enfermedades en niños y adultos. Las Hib causan meningitis bacteriana y otras infecciones invasivas en niños menores de 4 años, mientras que las NTHi se ha aislado de casos de otitis media, sinusitis, epiglotitis, traqueobronquitis y neumonía y pueden causar sepsis neonatal. Actualmente no existe una vacuna disponible comercialmente contra NTHi, pero se están utilizando varias vacunas contra Hib. Estas vacunas consisten en el polisacárido capsular de Hib, fosfato de polirribosil ribitol, conjugado con varios portadores de proteínas (complejo de membrana externa meningocócica, toxoide tetánico, toxina diftérica mutante no tóxica o toxoide diftérico) para superar la débil respuesta inmune al polisacárido capsular en niños menores de 18 meses de edad. Las proteínas de la membrana externa de *H. influenzae* (OMP) también se consideran portadores de fosfato de polirribosil ribitol, ya que se muestra que son objetivos de los anticuerpos del huésped después de las infecciones por Hib y NTHi. Los anticuerpos contra OMP P1, P2, P4, P5 y P6 y una proteína de 98 kDa han sido probados en protección *in vivo* y ensayos bactericidas *in vitro* contra infecciones por *H. influenzae*, con anticuerpos contra P1, P4 y P6 que muestran actividad biológica contra cepas de *H. influenzae* homólogas y heterólogas. La falta de protección heteróloga de los anticuerpos contra otras OMP se debe en parte a la diversidad antigénica de estas proteínas entre las diferentes cepas de *H. influenzae*. Por lo tanto, un antígeno ideal debe estar expuesto tanto en la superficie bacteriana como antigénicamente bien conservado. En este laboratorio, una proteína de membrana de 42 kDa (proteína D) que está ampliamente distribuida y conservada antigénicamente entre las cepas de Hib y NTHi se ha aislado, clonado, secuenciado y demostrado ser un factor de patogenicidad y un posible candidato a vacuna (1-5)

Hace dos décadas, se demostró que *Haemophilus influenzae* y *M. catarrhalis* mostraban una fuerte afinidad por la IgD humana soluble y unida a la superficie (6). La unión a la IgD parece ser paralela a una interacción similar con la IgD unida a la superficie a nivel celular, un fenómeno que explica los fuertes efectos mitogénicos en los linfocitos humanos por *H. influenzae* y *M. catarrhalis* (7-9). Se aisló y clonó una proteína de membrana externa de unión a IgD de *H. influenzae* (proteína D), y se demostró que era un factor de patogenicidad importante (1-5). Sin embargo, la proteína D no se une universalmente a todos los mielomas de IgD (10).

El documento WO 02/028889 divulga una secuencia, la SEQ ID NO: 10, que tiene un alto grado de identidad de aminoácidos con la SEQ ID NO: 1 de la presente solicitud. Describe los polipéptidos de *Haemophilus influenzae* útiles para prevenir o tratar las infecciones por *Haemophilus influenzae*, pero no proporciona datos que indiquen que la proteína E, o de hecho cualquiera de los polipéptidos descritos, pueda usarse con éxito en una vacuna.

Sumario de la invención

45 En vista del hecho de que se ha descubierto que *H. influenzae* es una causa principal de infecciones en las vías respiratorias superiores e inferiores, existe una necesidad actual de desarrollar vacunas que puedan usarse contra *H. influenzae*.

50 Por lo tanto, el objetivo de la presente invención ha sido descubrir de qué manera *H. influenzae* interactúa con las células del cuerpo e interactúa con el sistema inmunitario, y poder proporcionar un nuevo tipo de vacuna.

De acuerdo con un aspecto, la presente invención se refiere a una composición de la vacuna para uso en el tratamiento o prevención de otitis media, sinusitis, infecciones del tracto respiratorio inferior o enfermedad pulmonar obstructiva crónica, causadas por *Haemophilus influenzae*, que comprende una proteína expuesta en la superficie, que puede detectarse en *Haemophilus influenzae*, que tiene una secuencia de aminoácidos como se divulga en las SEQ ID NOs. 1, 2, 6, 7, 8 o 9.

60 En otro aspecto, la composición de la vacuna comprende un fragmento inmunogénico de la proteína expuesta en la superficie, cuyo fragmento puede detectarse en *Haemophilus influenzae*.

En otro aspecto, la composición de la vacuna comprende una proteína inmunogénica recombinante basada en la proteína, en la que los aminoácidos en las posiciones 1 a 21 de la SEQ ID NO: 1 se han eliminado o reemplazado por uno o más aminoácidos. En un aspecto, los aminoácidos en la posición 1-21 de la SEQ ID NO: 1 han sido reemplazados por una secuencia de 0 a 21 aminoácidos opcionales. En otro aspecto, la proteína tiene una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 2.

- En un aspecto, el fragmento tiene una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 6.
- En otro aspecto, el fragmento tiene una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 7.
- 5 De acuerdo con un aspecto, el fragmento tiene una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 8.
- En otro aspecto, el fragmento tiene una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 9.
- 10 De acuerdo con otro aspecto, la composición de la vacuna comprende al menos un dímero, trímero o multímero de la proteína, fragmento o péptido de acuerdo con lo anterior.
- En otro aspecto, la composición de la vacuna para el uso anterior comprende además uno o más adyuvantes, vehículos, excipientes, aglutinantes, portadores, conservantes, agentes tamponantes, agentes emulsionantes, agentes humectantes o compuestos facilitadores de la transfección farmacéuticamente aceptables.
- 15 De acuerdo con un aspecto adicional, la composición de la vacuna para el uso anterior comprende al menos una vacuna adicional.
- De acuerdo con otro aspecto, la composición de la vacuna para uso de acuerdo con lo anterior comprende una porción inmunogénica de otra molécula.
- 20 En un aspecto, la porción inmunogénica de otra molécula se elige del grupo que comprende Proteína D de *H. influenzae*, MID de *Moraxella catarrhalis*, UspA1 o UspA2 de *Moraxella catarrhalis* y proteína de membrana externa de cualquier patógeno del tracto respiratorio.
- 25 De acuerdo con otro aspecto, la composición de la vacuna para su uso de acuerdo con las reivindicaciones 1-3 adjuntas comprende un producto de fusión, en el que una proteína, fragmento o péptido de acuerdo con las reivindicaciones 1-5 adjuntas está covalentemente, o por cualquier otro medio, unido a una proteína, carbohidrato o matriz.
- 30 En otro aspecto, la composición de la vacuna comprende la secuencia de aminoácidos de la SEQ ID NO: 1.
- En otro aspecto, la presente invención proporciona una composición de la vacuna que comprende un polipéptido aislado de la SEQ ID NO: 1, como se define en la reivindicación 9 adjunta.
- 35 De acuerdo con un aspecto, en la composición de la vacuna, el polipéptido carece de un péptido señal (aminoácidos 1-20) de la SEQ ID NO: 1.
- En un aspecto, el polipéptido tiene una o más de las siguientes sustituciones:
- 40 Val en lugar de Ala en la posición 23 de la SEQ ID NO: 1,
Gln en lugar de Lys en la posición 24 de la SEQ ID NO: 1,
Met en lugar de Val en la posición 28 de la SEQ ID NO: 1,
Thr en lugar de Ala en la posición 31 de la SEQ ID NO: 1,
45 Val en lugar de Ile en la posición 41 de la SEQ ID NO: 1,
Ala en lugar de Val en la posición 47 de la SEQ ID NO: 1,
Lys en lugar de Arg en la posición 76 de la SEQ ID NO: 1,
Val en lugar de Ile en la posición 107 de la SEQ ID NO: 1,
Lys en lugar de Gly en la posición 152 de la SEQ ID NO: 1,
50 Val en lugar de Ala en la posición 154 de la SEQ ID NO: 1.
- De acuerdo con un aspecto, la composición de la vacuna comprende un polipéptido como se menciona en la reivindicación 9 adjunta, cuyo fragmento (si es necesario cuando está acoplado a un vehículo) es capaz de elevar una respuesta inmune que reconoce el polipéptido de la SEQ ID NO: 1 (o el polipéptido como se menciona en la reivindicación 9 adjunta respectivamente), o es capaz de unirse a la IgD humana.
- 55 En un aspecto, la composición de la vacuna comprende un polipéptido o un fragmento inmunogénico como se menciona en la reivindicación 9 adjunta, en el que dicho polipéptido o dicho fragmento inmunogénico es parte de una proteína de fusión más grande.
- 60 En un aspecto de la invención, la composición de la vacuna comprende una cantidad eficaz del polipéptido o el fragmento inmunogénico mencionado en cualquiera de las reivindicaciones 9 a 11 adjuntas y un excipiente farmacéuticamente aceptable.
- 65 En un aspecto, la composición de la vacuna comprende al menos otro antígeno de *Haemophilus influenzae*.

En otro aspecto, la composición de la vacuna mencionada en cualquiera de las reivindicaciones 1 a 13 adjuntas se formula con UspA1 y/o UspA2 de *Moraxella catarrhalis*.

5 En un aspecto, la composición de la vacuna mencionada en una cualquiera de las reivindicaciones 1 a 13 adjuntas se formula con la proteína D de *Haemophilus influenzae*.

En un aspecto, la *Haemophilus influenzae* está encapsulada o no es tipificable.

10 En otro aspecto, la composición de la vacuna es para la profilaxis o el tratamiento de otitis media, sinusitis o infecciones del tracto respiratorio inferior en niños y adultos que padecen, por ejemplo, enfermedad pulmonar obstructiva crónica (EPOC).

Descripción de las figuras

15 Figura 1. La proteína E de *Haemophilus influenzae* de 16,3 kDa es detectada por una proteína IgD(λ) de mieloma. En A, se muestra por análisis de citometría de flujo la expresión de pE en *H. influenzae* 772. Se muestra SDS-PAGE y transferencia Western (B) y análisis de electroforesis en gel de poliacrilamida SDS bidimensional (C) de proteínas de membrana externa tratadas con Empigen® de *H. influenzae* MinnA. Los extractos de proteínas de la membrana externa se muestran antes (B) y después de la separación en Q-Sepharose (C). La flecha en el panel C indica la localización predicha para pE basada en una transferencia Western (usando IgD(λ) como sonda) de un gel correspondiente. En A, las bacterias se cargaron con proteína IgD(λ) de mieloma seguido de incubación con pAb anti-IgD conjugado con FITC de conejo y análisis de citometría de flujo. En B, se muestran un gel SDS teñido con azul de Coomassie (mancha) y una transferencia Western sondeada con IgD(λ) de mieloma humano seguido por incubación con anticuerpos policlonales de cabra anti-IgD humana conjugados con peroxidasa de rábano picante. Las muestras se hirvieron en presencia de 2-mercaptoetanol durante 10 minutos antes de la carga.

20 Figura 2. Perfiles de citometría de flujo de *E. coli* que expresa pE en comparación con la *H. influenzae* 3655 tipo silvestre y un mutante deficiente en pE. La *E. coli* que alberga un vector pUC18 vacío (A) se compara con las bacterias transformadas con ADN genómico que contiene pUC18 de *H. influenzae* 772 (genes HI0175 a HI0178) (B). Se muestra la expresión de pE en la *H. influenzae* 3655 de tipo silvestre no tipificable (C) y el mutante correspondiente (D). La cepa de *E. coli* JM83 y *H. influenzae* se cultivaron durante la noche en cultivos líquidos. *E. coli* se incubó con IgD(λ) de mieloma humano en hielo. Después de 1 hora y lavados, se añadió pAb de conejo anti-IgD humana conjugado con FITC durante 30 minutos adicionales seguido de etapas de lavado y análisis de citometría de flujo posterior. El mismo procedimiento se realizó con *H. influenzae* 3655 o el mutante pE derivado usando anticuerpos policlonales de conejo anti-pE específicos y pAb de cabra anti-conejo conjugado con FITC.

30 Figura 3. Expresión de pE de *H. influenzae* y especies relacionadas como se revela por citometría de flujo y un suero de mieloma con IgD(λ). Se analizaron 22 cepas de NTHi y 27 cepas de especies de *Haemophilus* o bacterias relacionadas. Las bacterias crecieron hasta la fase estacionaria y se incubaron con IgD(λ) de mieloma humano en hielo. Después de 1 h y lavados, se agregaron anticuerpos policlonales (pAb) de conejo anti-IgD humana conjugados con FITC durante 30 minutos adicionales seguidos por etapas de lavado y análisis de citometría de flujo posteriores.

40 Figura 4. pE se expresa en NTHi y *H. influenzae* encapsulada como lo revelan las transferencias Western. Las proteínas bacterianas de las cepas indicadas se prepararon usando Empigen® y se sometieron a un gel SDS seguido de transferencia Western sondeadas con IgD(λ) de mieloma humano y pAb policlona de cabra anti-IgD humana conjugado con peroxidasa de rábano picante como anticuerpos de detección.

45 Figura 5. PE22-160 recombinante basada en la secuencia de NTHi 772 en comparación con pE nativa de *H. influenzae* MinnA. En A, la secuencia de terminal amino del fragmento A derivado de pE (SEQ ID NO: 32) en comparación con la secuencia predicha del terminal amino de la proteína nativa pE (SEQ ID NO: 31). En B, se muestra una ilustración esquemática de pE(A) con la etiqueta de Histidina (SEQ ID NO: 12) (MDIGINSDP divulgada como la SEQ ID NO: 33). En C, el tamaño y la pureza se demuestran en una PAGE teñida con Coomassie. En D y E, un extracto de proteína de membrana externa (OMP) de *H. influenzae* MinnA se compara con pE(A) producida de forma recombinante en un gel teñido con Coomassie y transferencia Western, respectivamente. En A, la secuencia del péptido señal se eliminó además del residuo de aminoácido glutamina 21. Se derivaron nueve aminoácidos del vector de expresión pET26(+) como se indica. Los números representan posiciones de aminoácidos que comienzan desde el inicio de la traducción de pE. La pE(A) recombinante se produjo en *E. coli*, se purificó y se sometió a degradación de Edman para analizar el sitio de escisión de la peptidasa señal. En D y E, se corrieron dos geles simultáneamente, uno se tiñó con azul brillante de Coomassie y otro se transfirió sobre membranas Immobilon-P, se sondeó con proteína de mieloma humano IgD(λ) seguido de incubación con anticuerpos secundarios apropiados conjugados con peroxidasa de rábano picante. La fracción de OMP se purificó usando Empigen® como se divulga en Materiales y métodos.

50 Figura 6. La proteína E se conserva extraordinariamente. Se muestra la frecuencia de mutaciones puntuales en 13 a 31 cepas de *Haemophilus influenzae* (Tabla 2), incluidos los aislados encapsulados y no tipificables. Los resultados se obtuvieron mediante secuenciación usando cebadores flanqueantes. Todas las secuencias se compararon con la secuencia de pE de *H. influenzae* Rd que se usó como secuencia de referencia y se muestra en el presente documento (la secuencia de ADN es la SEQ ID NO: 34 y la secuencia de proteína es la SEQ ID NO: 35).

60 Figura 7. El perfil de hidropatía de pE. Se indican las partes hidrofóbicas e hidrofílicas de los residuos de aminoácidos individuales. El péptido señal predicho también se divulga. Los datos se obtuvieron utilizando un método estándar como se divulga (21).

Figura 8. SDS-PAGE que muestra pE22-160 recombinante (fragmento A) y una serie de fragmentos truncados designados B a H. En A, se muestra un esquema de los diferentes fragmentos, mientras que en B, se muestra una SDS-PAGE. El ADN que codifica las diversas proteínas se ligó al vector de expresión pET26(+) y se expresó de forma recombinante en *E. coli*. Las proteínas sobreexpresadas resultantes se purificaron en resinas de níquel y se sometieron a separación en una SDS-PAGE seguido de tinción con azul brillante de Coomassie.

Figura 9. Una cepa mutante deficiente en pE (NTHi 3655) tiene una capacidad de 100 a 1.000 veces menor para inducir otitis media aguda en ratas. La infección se indujo en ratas Sprague-Dawley macho mediante una incisión ventral en la línea media del cuello, seguida por inyección en la cavidad del oído medio de los números indicados de bacterias en 0,05 mL. Los datos mostrados son del día 3 del desafío y son representativos de cinco animales en cada grupo.

Figura 10. Concentraciones medias de anticuerpos IgG e IgA dirigidos contra pE en sueros de diferentes grupos de edad. Los anticuerpos anti-pE se analizaron mediante un ELISA tipo sándwich utilizando pE(A) recombinante como cebo. La pureza de pE(A) fue como se indica en la Figura 5.

Figura 11. pE se conserva extraordinariamente dentro de diferentes cepas de *Haemophilus*. El gen pe fue secuenciado en *H. influenzae* encapsulada tipo a (n = 2), b (n = 2), c (n = 2), d (n = 1), e (n = 2) y f (n = 3), NTHi (n = 8), *H. influenzae* biovar *aegypticus* (n = 6) y *H. aegypticus* (n = 5), utilizando cebadores flanqueantes. Rd designa la cepa de *H. influenzae* Rd (Hi0178) y 772 la cepa 772 de NTHi. Los números 65 a 577 corresponden a las cepas descritas en la Tabla 1 (las 32 secuencias enumeradas se correlacionan con las SEQ ID NOs: 36, 1, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61 y 62, respectivamente, en orden de aparición de arriba hacia abajo).

Descripción de la invención

Debe entenderse que la fraseología y la terminología empleadas en este documento tienen fines descriptivos.

La presente solicitud divulga la clonación y expresión de una nueva proteína de membrana externa de *H. influenzae* denominada proteína E (pE). La proteína se descubrió usando un suero de mieloma con IgD(λ) humana con afinidad específica por pE.

Para maximizar el rendimiento de pE recombinante, se fabricó un fragmento de pE truncado que consiste en los residuos de aminoácidos lisina 22 a lisina 160. El péptido señal del terminal N que incluye el aminoácido glutamina 21 se eliminó de esta manera y se reemplazó con el péptido líder además de nueve residuos procedentes del vector pET26(+). La pE truncada (es decir, pE22-160) se designó pE(A).

La presente divulgación comprende la proteína pE de membrana externa de *Haemophilus* y los péptidos derivados de pE pE22-60, pE22-95, pE22-125, pE41-68, pE56-125, pE56-160, pE86-160, pE115-160 y dímeros, trímeros u oligómeros de los mismos. En particular, las secuencias de pE o los péptidos derivados que están expuestos en la superficie tienen mayor prioridad.

Por lo tanto, las composiciones de la vacuna de acuerdo con la presente divulgación comprenden como componentes inmunogénicos una proteína expuesta en la superficie, que puede detectarse en todas las *Haemophilus influenzae*, un fragmento inmunogénico de dicha proteína expuesta en la superficie, una proteína inmunogénica recombinante basada en dicha proteína expuesta en la superficie, una proteína inmunogénica recombinante que tiene una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 2, y/o un péptido que tiene una secuencia de aminoácidos de acuerdo con las SEQ ID NOs.: 4-10, o un fragmento de la misma. Las composiciones de la vacuna también pueden comprender una proteína o polipéptido de fusión, o un producto de fusión de acuerdo con la presente invención como componentes inmunogénicos. Los componentes inmunogénicos son capaces de provocar un anticuerpo u otra respuesta inmune a *Haemophilus influenzae*, en la que los anticuerpos inducidos inhiben la patogénesis de la bacteria *Haemophilus influenzae* en las células del sujeto. Una "dosis inmunogénica" de una composición de la vacuna de acuerdo con la invención es aquella que genera, después de la administración, una respuesta inmune humoral y/o celular detectable en comparación con una respuesta inmune estándar antes de la administración.

Las secuencias nucleicas usadas en las composiciones de la vacuna de la presente divulgación para generar los antígenos pueden insertarse en cualquiera de una amplia variedad de vectores de expresión mediante una variedad de procedimientos. Dichos procedimientos se consideran conocidos por los expertos en la materia.

Las composiciones de la vacuna se logran fácilmente usando métodos y técnicas bien conocidos, y se pueden administrar de varias maneras, preferiblemente de forma parenteral o intranasal. Las formulaciones adecuadas para administración parenteral o intranasal incluyen soluciones de inyección estériles acuosas y no acuosas que pueden contener antioxidantes, tampones, bacteriostáticos y solutos que hacen que la formulación sea isotónica con el fluido corporal del sujeto en cuestión; y suspensiones estériles acuosas y no acuosas que pueden incluir agentes de suspensión o agentes espesantes. El ingrediente inmunogénico activo a menudo se mezcla con excipientes que son farmacéuticamente aceptables, por ejemplo, agua, solución salina, dextrosa, glicerol, etanol o similares. Además, la composición de la vacuna también puede contener cantidades menores de sustancias auxiliares tales como agentes humectantes o emulsionantes, agentes tampón de pH, aglutinantes, vehículos o conservantes.

Las composiciones de la vacuna también pueden incluir adyuvantes para mejorar la inmunogenicidad de la composición, tales como adyuvantes de Freund y otros sistemas conocidos en la técnica. Los componentes inmunogénicos de las composiciones de la vacuna, es decir, las proteínas, fragmentos, péptidos, proteínas o polipéptidos de fusión, o productos de fusión de la presente invención, pueden formularse en la vacuna como formas neutras o salinas.

La dosificación de las composiciones de la vacuna dependerá de la actividad específica de la vacuna y puede determinarse fácilmente mediante experimentación de rutina. Las composiciones de la vacuna se administran en una cantidad que será terapéuticamente efectiva e inmunogénica, y la cantidad depende del sujeto.

La invención se refiere a polipéptidos y polinucleótidos de proteína E como se divulga con mayor detalle a continuación. En particular, la invención se refiere a polipéptidos y polinucleótidos de la proteína E de *H. Influenzae* no tipificable. Los polipéptidos de la proteína E tienen una secuencia señal y están expuestos en la superficie de la bacteria. El péptido señal se localiza desde el residuo 1 hasta el residuo 20 del polipéptido de la proteína E.

Una referencia a la "proteína E" en el presente documento es una referencia a cualquiera de los péptidos, fragmentos inmunogénicos, fusiones, polipéptidos o proteínas de la invención discutidos en el presente documento (tales como la SEQ ID NO: 1 con o sin la secuencia señal). Un "polinucleótido que codifica la proteína D" se refiere a cualquier secuencia de polinucleótidos que codifica cualquiera de los péptidos, fragmentos inmunogénicos, fusiones, polipéptidos o proteínas de la invención discutida en el presente documento.

El término "que comprende" en el presente documento alternativamente puede ser sustituido por el término "que consiste en".

La divulgación se refiere especialmente a polinucleótidos de proteína E y polipéptidos codificados enumerados en el presente documento.

Se entiende que las secuencias enumeradas en el Listado de secuencias a continuación como "ADN" representan un ejemplo de una realización de la invención, ya que los expertos en la técnica reconocerán que tales secuencias pueden emplearse de manera útil en polinucleótidos en general, incluyendo ribopolinucleótidos.

Las secuencias de los polinucleótidos de la proteína E se exponen en la SEQ ID NO: 11 (de la cepa 772 de NTHi). Las secuencias de los polipéptidos codificados de la proteína E se exponen en la SEQ ID NO: 1 (de la cepa 772 de NTHi), 2, 3, 4, 5, 6, 7, 8, 9, 10.

Polipéptidos

En un aspecto, se proporcionan polipéptidos de *H. influenzae* (en particular, *H. influenzae* no tipificable) a los que se hace referencia en el presente documento como "proteína E" y "polipéptidos de la proteína E", así como biológica, diagnóstica, profiláctica, clínica o terapéuticamente variantes de los mismos, útiles y composiciones que los comprenden

La presente divulgación proporciona además:

- (a) un polipéptido aislado que comprende una secuencia de aminoácidos que tiene al menos 85% de identidad, preferiblemente al menos 90% de identidad, más preferiblemente al menos 95% de identidad, lo más preferiblemente al menos 97-99% o identidad exacta, a la de cualquier secuencia de la SEQ ID NO: 1-2 y 4-10;
- (b) un polipéptido codificado por un polinucleótido aislado que comprende una secuencia de polinucleótidos que tiene al menos 85% de identidad, preferiblemente al menos 90% de identidad, más preferiblemente al menos 95% de identidad, incluso más preferiblemente al menos 97-99% o identidad exacta con cualquier secuencia de la SEQ ID NO: 11 en toda la longitud de la secuencia seleccionada de la SEQ ID NO: 11; o
- (c) un polipéptido codificado por un polinucleótido aislado que comprende una secuencia de polinucleótidos que codifica un polipéptido que tiene al menos 85% de identidad, preferiblemente al menos 90% de identidad, más preferiblemente al menos 95% de identidad, incluso más preferiblemente al menos 97-99% o identidad exacta, con la secuencia de aminoácidos de cualquier secuencia de la SEQ ID NO: 1-2 y 4-10.

Los polipéptidos de la proteína E proporcionados en las SEQ ID NOs: 1-2 y 4-10 son los polipéptidos de la proteína E de cepas de *H. influenzae* no tipificables. Se han determinado secuencias adicionales de proteína E a partir de cepas de *H. influenzae* enumeradas en la Tabla 1.

La divulgación también proporciona un fragmento inmunogénico de un polipéptido de la proteína E, es decir, una porción contigua del polipéptido de la proteína E que tiene la misma o sustancialmente la misma actividad inmunogénica que el polipéptido que comprende la secuencia de aminoácidos correspondiente seleccionada de las SEQ ID NOs: 1-2 y 4-10; es decir, el fragmento (si es necesario cuando se acopla a un vehículo) es capaz de generar una respuesta inmune que reconoce al polipéptido de la proteína E. Alternativamente, o además, el fragmento inmunogénico puede retener una función de unión a IgD de la proteína de longitud completa (como se divulga en la

sección de Ejemplos, por ejemplo, la capacidad de unir IgD(λ) del sitio de unión (Birmingham, Inglaterra). Dicho fragmento inmunogénico puede incluir, por ejemplo, el polipéptido de la proteína E que carece de una secuencia líder en el terminal N, y/o un dominio transmembrana y/o un dominio de anclaje en el terminal C. En un aspecto preferido, el fragmento inmunogénico de la proteína E comprende sustancialmente todo el dominio extracelular de un polipéptido que tiene al menos un 85% de identidad, preferiblemente al menos un 90% de identidad, más preferiblemente al menos un 95% de identidad, lo más preferiblemente al menos un 97-99% de identidad, con aquella secuencia seleccionada de las SEQ ID NOs: 1-2 y 4-10 en toda la longitud de dicha secuencia.

Un fragmento es un polipéptido que tiene una secuencia de aminoácidos que es completamente igual a una parte pero no a toda la secuencia de aminoácidos de cualquier polipéptido de la invención. Al igual que con los polipéptidos de la proteína E, los fragmentos pueden ser "independientes" o estar comprendidos dentro de un polipéptido más grande del cual forman una parte o región, lo más preferiblemente como una única región continua en un único polipéptido más grande. Por lo tanto, un fragmento puede ser más corto que la secuencia nativa de longitud completa o, si está comprendido dentro de un polipéptido más grande, puede ser una secuencia nativa de longitud completa o una proteína de fusión más larga.

Los fragmentos preferidos incluyen, por ejemplo, polipéptidos de truncamiento que tienen una porción de una secuencia de aminoácidos seleccionada de la SEQ ID NO: 1-2 y 4-10 o de variantes de la misma, como una serie continua de residuos que incluye una secuencia de aminoácidos del terminal amino y/o carboxilo. También se prefieren las formas de degradación de los polipéptidos de la invención producidos por o en una célula huésped. Más preferidos son los fragmentos caracterizados por atributos estructurales o funcionales tales como fragmentos que comprenden hélice alfa y regiones de formación de hélice alfa, lámina beta y regiones de formación de láminas beta, giros y regiones de formación de giros, espirales y regiones de formación de espirales, regiones hidrofílicas, regiones hidrófobas, regiones alfa anfipáticas, regiones beta anfipáticas, regiones flexibles, regiones formadoras de superficie, región de unión al sustrato y regiones de alto índice antigénico.

Los fragmentos preferidos adicionales incluyen un polipéptido aislado que comprende una secuencia de aminoácidos que tiene al menos 50 o 100 aminoácidos contiguos de una secuencia de aminoácidos seleccionada de las SEQ ID NOs: 1-2 y 4-10 o un polipéptido aislado que comprende una secuencia de aminoácidos que tiene al menos 50 o 100 aminoácidos contiguos truncados o eliminados de una secuencia de aminoácidos seleccionada de la SEQ ID NO: 1-2 y 4-10.

Fragmentos aún más preferidos son aquellos que comprenden un epítipo de células B, por ejemplo, aquellos fragmentos/péptidos descritos en el Ejemplo 10.

Se pueden emplear fragmentos de los polipéptidos de la invención para producir el polipéptido de longitud completa correspondiente mediante síntesis de péptidos; por lo tanto, estos fragmentos pueden emplearse como intermedios para producir los polipéptidos de longitud completa de la invención.

Se prefieren particularmente las variantes en las que varios, 5-10, 1-5, 1-3, 1-2 o 1 aminoácidos se sustituyen, eliminan o añaden en cualquier combinación.

Los polipéptidos, o fragmentos inmunogénicos, de la divulgación pueden estar en forma de la proteína "madura" o pueden ser parte de una proteína más grande tal como un precursor o una proteína de fusión. A menudo es ventajoso incluir una secuencia de aminoácidos adicional que contiene secuencias secretoras o líderes, prosecretoras, secuencias que ayudan en la purificación, tales como múltiples residuos de histidina, o una secuencia adicional para la estabilidad durante la producción recombinante. Además, también se considera la adición de polipéptidos exógenos o cola de lípidos o secuencias de polinucleótidos para aumentar el potencial inmunogénico de la molécula final.

En un aspecto, la divulgación se refiere a proteínas de fusión solubles modificadas genéticamente que comprenden un polipéptido de la presente invención, o un fragmento del mismo, y varias porciones de las regiones constantes de cadenas pesadas o ligeras de inmunoglobulinas de diversas subclases (IgG, IgM, IgA, IgE). Se prefiere como inmunoglobulina la parte constante de la cadena pesada de IgG humana, particularmente la IgG1, en la que la fusión tiene lugar en la región bisagra. En una realización particular, la parte Fc se puede eliminar simplemente mediante la incorporación de una secuencia de escisión que se puede escindir con el factor de coagulación sanguínea Xa.

Además, esta invención se refiere a procesos para la preparación de estas proteínas de fusión mediante ingeniería genética, y al uso de las mismas para la detección, diagnóstico y terapia de fármacos. Otro aspecto de la invención también se refiere a polinucleótidos que codifican tales proteínas de fusión. Se pueden encontrar ejemplos de tecnología de proteínas de fusión en las solicitudes de patente internacional Nos. WO94/29458 y WO94/22914.

Las proteínas pueden conjugarse químicamente o expresarse como proteínas de fusión recombinantes permitiendo que se produzcan niveles aumentados en un sistema de expresión en comparación con la proteína no fusionada. El compañero de fusión puede ayudar a proporcionar epítopos auxiliares T (compañero de fusión inmunológica), preferiblemente epítopos auxiliares T reconocidos por humanos, o ayudar a expresar la proteína (potenciador de la expresión) con rendimientos más altos que la proteína recombinante nativa. Preferiblemente, la pareja de fusión será tanto una pareja de fusión inmunológica como una pareja potenciadora de la expresión.

Los compañeros de fusión incluyen la proteína D de *Haemophilus influenzae* (EP 594610) y la proteína no estructural del virus de la influenza, NS1 (hemaglutinina). Otro compañero de fusión es la proteína conocida como Omp26 (documento WO 97/01638). Otro compañero de fusión es la proteína conocida como LytA. Preferiblemente se usa la porción del terminal C de la molécula. LytA se deriva de *Streptococcus pneumoniae* que sintetiza una N-acetil-L-alanina amidasa, amidasa LytA, (codificada por el gen *lytA* {Gene, 43 (1986) páginas 265-272}), una autolisina que degrada específicamente ciertos enlaces en la cadena principal del peptidoglicano. El dominio del terminal C de la proteína LytA es responsable de la afinidad con la colina o con algunos análogos de colina tales como DEAE. Esta propiedad ha sido explotada para el desarrollo de plásmidos que expresan C-LytA de *E. coli* útiles para la expresión de proteínas de fusión. La purificación de proteínas híbridas que contienen el fragmento C-LytA en su extremo amino ha sido descrita {*Biotechnology*: 10, (1992) páginas 795-798}. Es posible usar la porción repetida de la molécula LytA que se encuentra en el extremo terminal C comenzando en el residuo 178, por ejemplo los residuos 188-305.

La presente divulgación también incluye variantes de los polipéptidos mencionados anteriormente, es decir, polipéptidos que varían de los referentes por sustituciones conservadoras de aminoácidos, por lo que un residuo se sustituye por otro con características similares. Estas sustituciones típicas se encuentran entre Ala, Val, Leu e Ile; entre Ser y Thr; entre los residuos ácidos Asp y Glu; entre Asn y Gln; y entre los residuos básicos Lys y Arg; o residuos aromáticos Phe y Tyr.

Los polipéptidos de la presente divulgación se pueden preparar de cualquier manera adecuada. Dichos polipéptidos incluyen polipéptidos aislados de origen natural, polipéptidos producidos de forma recombinante, polipéptidos producidos sintéticamente o polipéptidos producidos mediante una combinación de estos métodos. Los medios para preparar tales polipéptidos se conocen bien en la técnica.

Se prefiere que un polipéptido de la divulgación se derive de *H. influenzae* no tipificable, sin embargo, se puede obtener preferiblemente de otros organismos del mismo género taxonómico. También se puede obtener un polipéptido de la invención, por ejemplo, de organismos de la misma familia u orden taxonómico.

Polinucleótidos

Es un objetivo de la divulgación proporcionar polinucleótidos que codifican polipéptidos de la proteína E, particularmente polinucleótidos que codifican los polipéptidos denominados en este documento proteína E.

En un ejemplo particularmente preferido, los polinucleótidos comprenden una región que codifica polipéptidos de Proteína E que comprenden secuencias expuestas en la SEQ ID NO: 11 que incluyen un gen de longitud completa, o una variante del mismo.

Los polinucleótidos de la proteína E proporcionados en la SEQ ID NO: 11 son los polinucleótidos de la proteína E de la cepa 772 de *H. influenzae* no tipificable. Se han determinado otras secuencias de genes que codifican la proteína E de las cepas de *H. influenzae* enumeradas en la Tabla 1.

Como un aspecto adicional de la divulgación, se proporcionan moléculas de ácido nucleico aisladas que codifican y/o expresan polipéptidos y polinucleótidos de proteína E, particularmente polipéptidos y polinucleótidos de proteína E de *H. influenzae* no tipificables, que incluyen, por ejemplo, ARN no procesados, ARN de ribozima, ARNm, ADNc, ADN genómico, ADN B y Z. Otras realizaciones de la divulgación incluyen polinucleótidos y polipéptidos biológica, diagnóstica, profiláctica, clínica o terapéuticamente útiles, y variantes de los mismos, y composiciones que comprenden los mismos.

Otro aspecto de la divulgación se refiere a polinucleótidos aislados, que incluyen al menos un gen de longitud completa, que codifica un polipéptido de proteína E que tiene una secuencia de aminoácidos deducida de la SEQ ID NO: 1-2 y 4-10 y polinucleótidos estrechamente relacionados con el mismo y variantes del mismo.

Otro ejemplo de la divulgación se refiere al polipéptido de la proteína E de *H. influenzae* no tipificable que comprende o que consiste en una secuencia de aminoácidos seleccionada de la SEQ ID NO: 1-2 y 4-10 o una variante del mismo.

Usando la información proporcionada en este documento, tal como las secuencias de polinucleótidos expuestas en la SEQ ID NO: 11, se puede obtener un polinucleótido que codifica polipéptidos de proteína E usando métodos de clonación y selección estándar, tales como los de clonación y secuenciación de fragmentos de ADN cromosómico de bacterias que usan células de la cepa 3224A (o 772) de *H. influenzae* no tipificable como material de partida, seguido por la obtención de un clon de longitud completa. Por ejemplo, para obtener una secuencia de polinucleótidos de la invención, tal como una secuencia de polinucleótidos dada en la SEQ ID NO: 11, típicamente una biblioteca de clones de ADN cromosómico de la cepa 3224A (o 772) de *H. influenzae* no tipificable en *E. coli* o algún otro huésped adecuado se sondea con un oligonucleótido radiomarcado, preferiblemente uno de 17-mer o más, derivado de una secuencia parcial. Los clones que portan ADN idéntico al de la sonda se pueden distinguir usando condiciones de hibridación estrictas. Al secuenciar los clones individuales así identificados por hibridación con cebadores de secuenciación diseñados a partir del polipéptido original o secuencia de polinucleótidos, es posible extender la secuencia de

- polinucleótidos en ambas direcciones para determinar una secuencia génica de longitud completa. Convenientemente, dicha secuenciación se lleva a cabo, por ejemplo, usando ADN desnaturalizado de doble cadena preparado a partir de un clon de plásmido. Las técnicas adecuadas son descritas por Maniatis, T., Fritsch, E.F. y Sambrook et al., MOLECULAR CLONING, A LABORATORY MANUAL, 2nd Ed.; divulgan técnicas adecuadas. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, Nueva York (1989). (véase en particular Screening By Hybridization 1.90 and Sequencing Denatured Double-Stranded DNA Templates 13.70). La secuenciación directa del ADN genómico también se puede realizar para obtener una secuencia génica de longitud completa. Ilustrativo de la invención, el polinucleótido expuesto en la SEQ ID NO: 11 se descubrió en una biblioteca de ADN derivada de *H. influenzae* no tipificable.
- 5
- 10 Además, cada secuencia de ADN expuesta en la SEQ ID NO: 11 contiene un marco de lectura abierto que codifica una proteína que tiene aproximadamente el número de residuos de aminoácidos expuestos en la SEQ ID NO: 1 con un peso molecular deducido que puede calcularse usando valores de peso molecular de residuos de aminoácidos bien conocidos por los expertos en la técnica.
- 15 Los polinucleótidos de la SEQ ID NO: 11, entre el codón de inicio y el codón de parada, codifican respectivamente el polipéptido de la SEQ ID NO: 1.
- En otro aspecto, la presente divulgación proporciona un polinucleótido aislado que comprende o que consiste en:
- 20 (a) una secuencia de polinucleótidos que tiene al menos 85% de identidad, preferiblemente al menos 90% de identidad, más preferiblemente al menos 95% de identidad, incluso más preferiblemente al menos 97-99% o identidad exacta, con cualquier secuencia de polinucleótidos de la SEQ ID NO: 11 en toda la longitud de la secuencia de polinucleótidos de la SEQ ID NO: 11; o
- 25 (b) una secuencia de polinucleótidos que codifica un polipéptido que tiene al menos 85% de identidad, preferiblemente al menos 90% de identidad, más preferiblemente al menos 95% de identidad, incluso más preferiblemente al menos 97-99% o 100% de identidad exacta, con cualquier secuencia de aminoácidos seleccionada de la SEQ ID NO: 1-2 y 4-10 (o fragmento de la misma), a lo largo de toda la longitud de la secuencia de aminoácidos de la SEQ ID NO: 1-2 y 4-10 (o fragmento).
- 30 Un polinucleótido que codifica un polipéptido de la presente divulgación que incluye homólogos y ortólogos de especies distintas de *H. influenzae* no tipificable, puede obtenerse mediante un proceso que comprende las etapas de seleccionar una biblioteca apropiada en condiciones de hibridación rigurosas (por ejemplo, usando una temperatura en el intervalo de 45-65 °C y una concentración de SDS de 0,1-1%) con una sonda marcada o detectable que consiste o comprende cualquier secuencia seleccionada de la SEQ ID NO: 11 o un fragmento de la misma; y aislar un gen de
- 35 longitud completa y/o clones genómicos que contienen dicha secuencia de polinucleótidos.
- La divulgación proporciona una secuencia de polinucleótidos idéntica en toda su longitud a una secuencia de codificación (marco de lectura abierto) expuesta en la SEQ ID NO: 11. También se proporciona una secuencia de codificación para un polipéptido maduro o un fragmento del mismo, por sí mismo así como una secuencia de
- 40 codificación para un polipéptido maduro o un fragmento en el marco de lectura con otra secuencia de codificación, tal como una secuencia que codifica una secuencia líder o secretora, una secuencia preproteína, o proproteína o preproproteína. El polinucleótido de la divulgación también puede contener al menos una secuencia no codificante, que incluye, por ejemplo, pero no se limita a, al menos una secuencia 5' y 3' no codificante, tal como las secuencias transcritas pero no traducidas, señales de terminación (tales como señales de terminación dependientes de rho e independientes de rho), sitios de unión a ribosomas, secuencias de Kozak, secuencias que estabilizan ARNm, intrones y señales de poliadenilación. La secuencia polinucleotídica también puede comprender una secuencia codificante adicional que codifica aminoácidos adicionales. Por ejemplo, se puede codificar una secuencia marcadora que facilite la purificación del polipéptido fusionado. En ciertas realizaciones de la invención, la secuencia marcadora es un péptido de hexahistidina (SEQ ID NO: 12), como se proporciona en el vector pQE (Qiagen, Inc.) y se divulga en Gentz et al., Proc. Natl. Acad. Sci., USA 86: 821-824 (1989), o una etiqueta de péptido HA (Wilson et al., Cell 37: 767 (1984)), los cuales pueden ser útiles para purificar la secuencia de polipéptidos fusionados a ellos. Polinucleótidos de la divulgación también incluyen, pero no se limitan a, polinucleótidos que comprenden un gen estructural y sus secuencias asociadas naturalmente que controlan la expresión génica.
- 45
- 50 La secuencia de nucleótidos que codifica el polipéptido de la proteína E de la SEQ ID NO: 1-2 y 4-10 puede ser idéntica a la secuencia correspondiente que codifica el polinucleótido de la SEQ ID NO: 11 (o comprendida dentro de la SEQ ID NO: 11). Alternativamente, puede ser cualquier secuencia que, como resultado de la redundancia (degeneración) del código genético, también codifique un polipéptido de la SEQ ID NO: 1-2 y 4-10.
- 55
- 60 El término "polinucleótido que codifica un polipéptido" como se usa en el presente documento abarca polinucleótidos que incluyen una secuencia que codifica un polipéptido de la divulgación, particularmente un polipéptido bacteriano y más particularmente un polipéptido de la proteína E de *H. influenzae* no tipificable que tiene una secuencia de aminoácidos expuesta en cualquiera de las secuencias de la SEQ ID NO: 1-2 y 4-10 o fragmentos de las mismas. El término también abarca polinucleótidos que incluyen una única región continua o regiones discontinuas que codifican el polipéptido (por ejemplo, polinucleótidos interrumpidos por un fago integrado, una secuencia de inserción integrada, una secuencia de vector integrada, una secuencia de transposón integrada, o debido a la edición del ARN o la
- 65

reorganización del ADN genómico) junto con regiones adicionales, que también pueden contener secuencias codificantes y/o no codificantes.

5 La divulgación se refiere además a variantes de los polinucleótidos descritos en el presente documento que codifican variantes de un polipéptido que tiene una secuencia de aminoácidos deducida de cualquiera de las secuencias de la SEQ ID NO: 1-2 y 4-10. Se pueden usar fragmentos de polinucleótidos, por ejemplo, para sintetizar polinucleótidos de longitud completa de la divulgación.

10 Los fragmentos preferidos son aquellos polinucleótidos que codifican un epítipo de células B, por ejemplo, los fragmentos/péptidos descritos en el Ejemplo 10, y genes quiméricos recombinantes que comprenden dichos fragmentos de polinucleótidos.

15 Otras realizaciones particularmente preferidas son polinucleótidos que codifican variantes de proteína E, que tienen la secuencia de aminoácidos del polipéptido de proteína E de cualquier secuencia de la SEQ ID NO: 1-2 y 4-10 en la que varios, algunos, 5 a 10, 1 a 5, 1 a 3, 2, 1 o ningún residuo de aminoácido se sustituye, modifica, elimina o agrega en cualquier combinación. Especialmente preferidos entre estos son las sustituciones, adiciones y eliminaciones silenciosas, que no alteran las propiedades y actividades del polipéptido de la proteína E (por ejemplo, las propiedades descritas en la sección de ejemplos en este documento).

20 Otras realizaciones preferidas de la divulgación son polinucleótidos que son al menos 85% idénticos en toda su longitud a polinucleótidos que codifican polipéptidos de proteína E que tienen una secuencia de aminoácidos expuesta en cualquiera de las secuencias de la SEQ ID NO: 1-2 y 4 -10, y polinucleótidos que son complementarios a dichos polinucleótidos. Alternativamente, los más preferidos son los polinucleótidos que comprenden una región que es al menos 90% idéntica en toda su longitud a los polinucleótidos que codifican polipéptidos de proteína E y polinucleótidos complementarios a los mismos. A este respecto, se prefieren particularmente los polinucleótidos al menos 95% idénticos en toda su longitud a los mismos. Además, aquellos con al menos 97% son altamente preferidos entre aquellos con al menos 95%, y entre estos aquellos con al menos 98% y al menos 99% son particularmente altamente preferidos, siendo al menos 99% los más preferidos.

30 Las realizaciones preferidas son polinucleótidos que codifican polipéptidos que retienen sustancialmente la misma función o actividad biológica que el polipéptido maduro codificado por una secuencia de ADN seleccionada de la SEQ ID NO: 11 (por ejemplo, las actividades descritas en la sección de Ejemplos en el presente documento).

35 De acuerdo con ciertas realizaciones preferidas de esta divulgación, se proporcionan polinucleótidos que hibridan, particularmente en condiciones rigurosas, con secuencias de polinucleótidos de proteína E, tales como aquellos polinucleótidos de la SEQ ID NO: 11.

40 La divulgación se refiere además a polinucleótidos que hibridan con las secuencias de polinucleótidos proporcionadas en el presente documento. A este respecto, la invención se refiere especialmente a polinucleótidos que hibridan en condiciones rigurosas con los polinucleótidos descritos en el presente documento. Como se usa en el presente documento, los términos "condiciones rigurosas" y "condiciones de hibridación rigurosas" significan que la hibridación se produce solo si hay al menos un 95% y preferiblemente al menos un 97% de identidad entre las secuencias. Un ejemplo específico de condiciones de hibridación rigurosas es la incubación durante la noche a 42 °C en una solución que comprende: 50% de formamida, SSC 5x (NaCl 150mM, citrato trisódico 15mM), fosfato de sodio 50 mM (pH 7,6), solución 5x de Denhardt, 10% sulfato de dextrano y 20 microgramos/mL de ADN de esperma de salmón cortado y desnaturalizado, seguido de lavado del soporte de hibridación en SSC 0,1x a aproximadamente 65 °C. Las condiciones de hibridación y lavado son bien conocidas y ejemplificadas en Sambrook, et al., Molecular Cloning: A Laboratory Manual, segunda edición, Cold Spring Harbor, N.Y., (1989), particularmente el Capítulo 11. La hibridación en solución también se puede usar con las secuencias de polinucleótidos proporcionadas por la invención.

50 La divulgación también proporciona un polinucleótido que consiste en o que comprende una secuencia de polinucleótidos obtenida mediante la selección de una biblioteca apropiada que contiene el gen completo para una secuencia de polinucleótidos expuesta en cualquiera de las secuencias de la SEQ ID NO: 11 bajo condiciones de hibridación rigurosas con una sonda que tiene la secuencia de dicha secuencia de polinucleótidos expuesta en la secuencia correspondiente de la SEQ ID NO: 11 o un fragmento de la misma; y aislar dicha secuencia de polinucleótidos. Los fragmentos útiles para obtener dicho polinucleótido incluyen, por ejemplo, sondas y cebadores completamente descritos en otra parte del presente documento.

60 Como se discute en otra parte en este documento con respecto a los ensayos de polinucleótidos de la invención, por ejemplo, los polinucleótidos de la invención, se pueden usar como una sonda de hibridación para ARN, ADNc y ADN genómico para aislar ADNc de longitud completa y clones genómicos que codifican proteína E y para aislar ADNc y clones genómicos de otros genes que tienen una identidad alta, particularmente una identidad de secuencia alta, con los genes de la proteína E. Tales sondas generalmente comprenderán al menos 15 residuos de nucleótidos o pares de bases. Preferiblemente, tales sondas tendrán al menos 30 residuos de nucleótidos o pares de bases y pueden tener al menos 50 residuos de nucleótidos o pares de bases. Las sondas particularmente preferidas tendrán al menos 20 residuos de nucleótidos o pares de bases y tendrán menos de 30 residuos de nucleótidos o pares de bases.

Una región codificante de los genes de la proteína E puede aislarse mediante selección usando una secuencia de ADN proporcionada en la SEQ ID NO: 11 para sintetizar una sonda de oligonucleótidos. Luego se usa un oligonucleótido marcado que tiene una secuencia complementaria a la de un gen de la divulgación para seleccionar una biblioteca de ADNc, ADN genómico o ARNm para determinar con qué miembros de la biblioteca se hibrida la sonda.

Existen varios métodos disponibles y bien conocidos por los expertos en la técnica para obtener ADN de longitud completa, o extender ADN cortos, por ejemplo, aquellos basados en el método de amplificación rápida de extremos de ADNc (RACE) (véase, por ejemplo, Frohman, et al., PNAS USA 85: 8998-9002, 1988). Las modificaciones recientes de la técnica, ejemplificadas por la tecnología Marathon^{MR} (Clontech Laboratories Inc.), por ejemplo, han simplificado significativamente la búsqueda de ADNc más largos. En la tecnología Marathon^{MR}, los ADNc se han preparado a partir de ARNm extraído de un tejido elegido y una secuencia 'adaptadora' ligada en cada extremo. La amplificación de ácido nucleico (PCR) se lleva a cabo para amplificar el extremo 5' "faltante" del ADN usando una combinación de cebadores oligonucleotídicos específicos de gen y específicos de adaptador. Luego, la reacción de PCR se repite usando cebadores "anidados", es decir, cebadores diseñados para hibridar dentro del producto amplificado (típicamente un cebador específico del adaptador que hibrida además 3' en la secuencia del adaptador y un cebador específico del gen que hibrida además 5' en la secuencia génica seleccionada). Los productos de esta reacción se pueden analizar mediante secuenciación de ADN y construir un ADN de longitud completa, ya sea uniendo el producto directamente al ADN existente para obtener una secuencia completa, o llevando a cabo una PCR de longitud completa separada utilizando la nueva información de secuencia para el diseño del cebador 5'.

Los polinucleótidos y polipéptidos pueden emplearse, por ejemplo, como reactivos de investigación y materiales para el descubrimiento de tratamientos y diagnósticos de enfermedades, particularmente enfermedades humanas, como se divulga adicionalmente en este documento en relación con ensayos de polinucleótidos.

Los polinucleótidos que son oligonucleótidos derivados de una secuencia de la SEQ ID NO: 11 se pueden usar en los procesos descritos en el presente documento, pero preferiblemente para PCR, para determinar si los polinucleótidos identificados en el presente documento en su totalidad o en parte se transcriben en bacterias en el tejido infectado. Se reconoce que tales secuencias también tendrán utilidad en el diagnóstico de la etapa de infección y el tipo de infección que ha alcanzado el patógeno.

La divulgación también proporciona polinucleótidos que codifican un polipéptido que es la proteína madura más aminoácidos adicionales del terminal amino o carboxilo, o aminoácidos interiores del polipéptido maduro (cuando la forma madura tiene más de una cadena de polipéptidos, por ejemplo). Dichas secuencias pueden desempeñar un papel en el procesamiento de una proteína desde el precursor hasta una forma madura, pueden permitir el transporte de proteínas, pueden alargar o acortar la vida media de la proteína o pueden facilitar la manipulación de una proteína para su análisis o producción, entre otras cosas. Como generalmente es el caso *in vivo*, los aminoácidos adicionales pueden procesarse lejos de la proteína madura mediante enzimas celulares.

Para todos y cada uno de los polinucleótidos de la divulgación, se proporciona un polinucleótido complementario al mismo. Se prefiere que estos polinucleótidos complementarios sean completamente complementarios a cada polinucleótido con el que son complementarios.

Una proteína precursora, que tiene una forma madura del polipéptido fusionado a una o más prosequencias puede ser una forma inactiva del polipéptido. Cuando se eliminan las prosequencias, tales precursores inactivos generalmente se activan. Algunas o todas las prosequencias pueden eliminarse antes de la activación. En general, dichos precursores se llaman proproteínas.

Además de las representaciones estándar de A, G, C, T/U para nucleótidos, el término "N" también puede usarse para describir ciertos polinucleótidos de la invención. "N" significa que cualquiera de los cuatro nucleótidos de ADN o ARN puede aparecer en dicha posición designada en la secuencia de ADN o ARN, excepto que se prefiere que N no sea un ácido nucleico que cuando se toma en combinación con posiciones de nucleótidos adyacentes, cuando se lee en el marco de lectura correcto, tendría el efecto de generar un codón de terminación prematuro en dicho marco de lectura.

En resumen, un polinucleótido de la divulgación puede codificar una proteína madura, una proteína madura más una secuencia líder (que puede denominarse preproteína), un precursor de una proteína madura que tiene una o más prosequencias que no son secuencias líder de una preproteína, o una preproteína, que es un precursor de una proproteína, que tiene una secuencia líder y una o más prosequencias, que generalmente se eliminan durante las etapas de procesamiento que producen formas activas y maduras del polipéptido.

De acuerdo con un aspecto de la divulgación, se proporciona el uso de un polinucleótido de la invención con fines terapéuticos o profilácticos, en particular inmunización genética.

El uso de un polinucleótido de la invención en inmunización genética empleará preferiblemente un método de administración adecuado tal como inyección directa de ADN plasmídico en los músculos (Wolff et al., Hum Mol Genet (1992) 1: 363, Manthorpe et al., Hum. Gene Ther. (1983) 4: 419), suministro de ADN complejo con portadores de proteínas específicos (Wu et al., J Biol Chem. (1989) 264: 16985), coprecipitación de ADN con fosfato de calcio (Benvenisty & Reshef, PNAS USA, (1986) 83: 9551), encapsulación de ADN en diversas formas de liposomas (Kaneda et al., Science (1989) 243: 375), bombardeo de partículas (Tang et al., Nature (1992) 356: 152, Eisenbraun et al., DNA Cell Biol (1993) 12: 791) e infección *in vivo* usando vectores retrovirales clonados (Seeger et al., PNAS USA (1984) 81: 5849).

10 Vectores, células huésped, sistemas de expresión

La divulgación también se refiere a vectores que comprenden un polinucleótido o polinucleótidos de la divulgación, células huésped que están genéticamente modificadas con vectores de la divulgación y la producción de polipéptidos de la invención mediante técnicas recombinantes. Los sistemas de traducción libres de células también pueden emplearse para producir tales proteínas usando ARN derivados de los constructos de ADN de la divulgación.

15 Los polipéptidos recombinantes pueden prepararse mediante procesos bien conocidos por los expertos en la materia a partir de células huésped genéticamente modificadas que comprenden sistemas de expresión. La presente divulgación también divulga sistemas de expresión que comprenden un polinucleótido o polinucleótidos para células huésped que están genéticamente modificadas con tales sistemas de expresión.

20 Para la producción recombinante de los polipéptidos de la divulgación, las células huésped pueden modificarse genéticamente para incorporar sistemas de expresión o porciones de los mismos o polinucleótidos de la divulgación. La introducción de un polinucleótido en la célula huésped se puede efectuar mediante métodos descritos en muchos manuales de laboratorio estándar, tal como Davis, et al., BASIC METHODS IN MOLECULAR BIOLOGY, (1986) y Sambrook, et al., MOLECULAR CLONING: A LABORATORY MANUAL, segunda Edición, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989), tales como transfección con fosfato de calcio, transfección mediada por DEAE-dextrano, transvección, microinyección, transfección mediada por lípidos catiónicos, electroporación, conjugación, transducción, raspado de carga, introducción balística e infección.

25 Los ejemplos representativos de huéspedes apropiados incluyen células bacterianas, tales como células de estreptococos, estafilococos, enterococos, *E. coli*, estreptomicos, cianobacterias, *Bacillus subtilis*, *Neisseria meningitidis*, *Haemophilus influenzae* y *Moraxella catarrhalis*; células fúngicas, tales como células de levadura, *Kluyveromyces*, *Saccharomyces*, *Pichia*, un basidiomiceto, *Candida albicans* y *Aspergillus*; células de insecto tales como células de *Drosophila S2* y *Spodoptera Sf9*; células animales tales como células de melanoma CHO, COS, HeLa, C127, 3T3, BHK, 293, CV-1 y de Bowes; y células vegetales, tales como las células de una gimnosperma o angiosperma.

30 Se puede usar una gran variedad de sistemas de expresión para producir los polipéptidos de la divulgación. Dichos vectores incluyen, entre otros, vectores derivados de cromosomas, episomas y virus, por ejemplo, vectores derivados de plásmidos bacterianos, de bacteriófagos, de transposones, de episomas de levadura, de elementos de inserción, de elementos cromosómicos de levadura, de virus tales como baculovirus, papovavirus, tales como SV40, virus vaccinia, adenovirus, virus de la viruela de las aves, virus de la pseudorrabia, picornavirus, retrovirus y alfavirus y vectores derivados de combinaciones de los mismos, tales como los derivados de elementos genéticos de plásmidos y bacteriófagos, tales como los cósmidos y los fagémidos. Los constructos del sistema de expresión pueden contener regiones de control que regulan así como engendran expresión. Generalmente, cualquier sistema o vector adecuado para mantener, propagar o expresar polinucleótidos y/o expresar un polipéptido en un huésped puede usarse para la expresión a este respecto. La secuencia de ADN apropiada puede insertarse en el sistema de expresión mediante cualquiera de una variedad de técnicas conocidas y rutinarias, tales como, por ejemplo, las expuestas en Sambrook et al., MOLECULAR CLONING, A LABORATORY MANUAL, (citado anteriormente).

35 En sistemas de expresión recombinante en eucariotas, para la secreción de una proteína traducida en el lumen del retículo endoplásmico, en el espacio periplásmico o en el entorno extracelular, se pueden incorporar señales de secreción apropiadas en el polipéptido expresado. Estas señales pueden ser endógenas al polipéptido o pueden ser señales heterólogas.

40 Los polipéptidos de la presente divulgación pueden recuperarse y purificarse de cultivos celulares recombinantes mediante métodos bien conocidos que incluyen precipitación con sulfato de amonio o etanol, extracción con ácido, cromatografía de intercambio aniónico o catiónico, cromatografía de fosfocelulosa, cromatografía de interacción hidrófoba, cromatografía de afinidad, cromatografía de hidroxilapatita y cromatografía de lectina. Lo más preferiblemente, la cromatografía de afinidad de iones metálicos (IMAC) se emplea para la purificación. Se pueden emplear técnicas bien conocidas para desplegar proteínas para regenerar la conformación activa cuando el polipéptido se desnaturaliza durante la síntesis, el aislamiento y/o purificación intracelular.

45 El sistema de expresión también puede ser un microorganismo vivo recombinante, tal como un virus o una bacteria. El gen de interés puede insertarse en el genoma de un virus o bacteria recombinante vivo. La inoculación y la infección

in vivo con este vector vivo conducirán a la expresión *in vivo* del antígeno y a la inducción de respuestas inmunes. Los virus y bacterias utilizados para este propósito son, por ejemplo: virus de la viruela (por ejemplo, vaccinia, viruela aviar, viruela del canario), alfavirus (virus Sindbis, virus del bosque Semliki, virus de la encefalitis equina venezolana), adenovirus, virus adenoasociados, picornavirus (virus de la polio, rinovirus), virus del herpes (virus de la varicela zoster, etc.), listeria, salmonella, Shigella, BCG, estreptococos. Estos virus y bacterias pueden ser virulentos o atenuados de varias maneras para obtener vacunas vivas. Dichas vacunas vivas también forman parte de la divulgación.

Ensayos de diagnóstico, pronóstico, serotipado y mutaciones

Esta divulgación también está relacionada con el uso de polinucleótidos y polipéptidos de la proteína E de la invención para su uso como reactivos de diagnóstico. La detección de polinucleótidos y/o polipéptidos de la proteína E en un eucariota, particularmente un mamífero, y especialmente un ser humano, proporcionará un método de diagnóstico para el diagnóstico de enfermedad, estadificación de la enfermedad o respuesta de un organismo infeccioso a los fármacos. Los eucariotas, particularmente los mamíferos, y especialmente los humanos, particularmente aquellos infectados o sospechosos de estar infectados con un organismo que comprende los genes o proteínas de la proteína E, pueden detectarse a nivel de ácido nucleico o de aminoácidos mediante una variedad de técnicas bien conocidas, así como mediante métodos proporcionados en este documento.

Los polipéptidos y polinucleótidos para el pronóstico, diagnóstico u otro análisis pueden obtenerse a partir de materiales corporales de un individuo infectado y/o infectado putativamente. Los polinucleótidos de cualquiera de estas fuentes, particularmente ADN o ARN, pueden usarse directamente para la detección o pueden amplificarse enzimáticamente usando PCR o cualquier otra técnica de amplificación antes del análisis. El ARN, particularmente el ARNm, el ADNc y el ADN genómico también se pueden usar de la misma manera. Mediante la amplificación, se puede realizar la caracterización de la especie y la cepa del organismo infeccioso o residente presente en un individuo, mediante un análisis del genotipo de un polinucleótido seleccionado del organismo. Las eliminaciones e inserciones pueden detectarse mediante un cambio en el tamaño del producto amplificado en comparación con un genotipo de una secuencia de referencia seleccionada de un organismo relacionado, preferiblemente una especie diferente del mismo género o una cepa diferente de la misma especie. Las mutaciones puntuales se pueden identificar hibridando ADN amplificado con secuencias de polinucleótidos de proteína E marcadas. Las secuencias perfectamente o significativamente emparejadas pueden distinguirse de los dúplex imperfectamente o más significativamente no emparejados por digestión con DNasa o RNasa, para ADN o ARN respectivamente, o mediante la detección de diferencias en las temperaturas de fusión o la cinética de renaturalización. Las diferencias en la secuencia de polinucleótidos también pueden detectarse mediante alteraciones en la movilidad electroforética de fragmentos de polinucleótidos en geles en comparación con una secuencia de referencia. Esto puede llevarse a cabo con o sin agentes desnaturalizantes. Las diferencias de polinucleótidos también pueden detectarse mediante secuenciación directa de ADN o ARN. Véase, por ejemplo, Myers et al., Science, 230: 1242 (1985). Los cambios de secuencia en ubicaciones específicas también pueden revelarse mediante ensayos de protección de nucleasa, tales como RNasa, ensayos de protección, V1 y S1 o un método de escisión química. Véase, por ejemplo, Cotton et al., Proc. Natl. Acad. Sci., USA, 85: 4397-4401 (1985).

También se divulga un arreglo de sondas de oligonucleótidos que comprenden secuencias de nucleótidos de proteína E o fragmentos de las mismas que se pueden construir para realizar una selección eficiente de, por ejemplo, mutaciones genéticas, serotipos, clasificación o identificación taxonómica. Los métodos de tecnología de matriz son bien conocidos y tienen aplicabilidad general y se pueden usar para abordar una variedad de preguntas en genética molecular, incluida la expresión génica, el enlace genético y la variabilidad genética (véase por ejemplo, Chee et al., Science, 274: 610 (1996)).

También se divulga un kit de diagnóstico que comprende:

- (a) un polinucleótido de la presente invención, preferiblemente cualquiera de las secuencias de nucleótidos de la SEQ ID NO: 11, o un fragmento de la misma;
- (b) una secuencia de nucleótidos complementaria a la de (a);
- (c) un polipéptido de la presente invención, preferiblemente cualquiera de los polipéptidos de la SEQ ID NO: 1-10 o un fragmento del mismo; o
- (d) un anticuerpo contra un polipéptido de la presente invención, preferiblemente contra cualquiera de los polipéptidos de la SEQ ID NO: 1-10.

Se apreciará que en cualquiera de tales kits, (a), (b), (c) o (d) pueden comprender un componente sustancial. Tal kit será útil para diagnosticar una enfermedad o susceptibilidad a una enfermedad, entre otros.

También se divulga el uso de polinucleótidos como reactivos de diagnóstico. La detección de una forma mutada de un polinucleótido de la invención, preferiblemente cualquier secuencia de la SEQ ID NO: 11, que se asocia con una enfermedad o patogenicidad proporcionará una herramienta de diagnóstico que puede añadirse a, o definir, un diagnóstico de una enfermedad, un pronóstico del curso de la enfermedad, una determinación de un estadio de la enfermedad, o una susceptibilidad a una enfermedad, que resulta de la subexpresión, sobreexpresión o expresión alterada del polinucleótido. Los organismos, particularmente organismos infecciosos, que portan mutaciones en tal

polinucleótido se pueden detectar a nivel del polinucleótido mediante una variedad de técnicas, tales como las descritas en otra parte en este documento.

Las células de un organismo portador de mutaciones o polimorfismos (variaciones alélicas) en un polinucleótido y/o polipéptido de la invención también se pueden detectar a nivel de polinucleótido o polipéptido mediante una variedad de técnicas, para permitir una serotipificación, por ejemplo. Por ejemplo, RT-PCR se puede usar para detectar mutaciones en el ARN. Se prefiere particularmente usar RT-PCR junto con sistemas de detección automatizados, tales como, por ejemplo, GeneScan. ARN, ADNc o ADN genómico también se pueden usar para el mismo propósito, PCR. Como ejemplo, cebadores de PCR complementarios a un polinucleótido que codifican polipéptidos de la proteína E se pueden utilizar para identificar y analizar mutaciones.

También se divulgan cebadores con 1, 2, 3 o 4 nucleótidos eliminados del extremo 5' y/o el extremo 3'. Estos cebadores pueden usarse para, entre otras cosas, amplificar ADN y/o ARN de la proteína E aislada a partir de una muestra derivada de un individuo, tal como un material corporal. Los cebadores pueden usarse para amplificar un polinucleótido aislado de un individuo infectado, tal que el polinucleótido puede someterse entonces a diversas técnicas para elucidación de la secuencia de polinucleótido. De esta manera, las mutaciones en la secuencia de polinucleótidos se pueden detectar y usar para diagnosticar y/o pronosticar la infección o su estadio o transcurso, o para el serotipo y/o clasificar el agente infeccioso.

También se divulga un proceso para diagnosticar, enfermedad, preferiblemente infecciones bacterianas, más preferiblemente infecciones causadas por *H. influenzae* no tipificable, que comprende determinar a partir de una muestra derivada de un individuo, tal como un material corporal, un nivel aumentado de expresión de polinucleótido que tiene una secuencia de cualquiera de las secuencias de la SEQ ID NO: 11. La expresión aumentada o disminuida del polinucleótido de la proteína E se puede medir usando cualquiera de los métodos bien conocidos en la técnica para la cuantificación de polinucleótidos, tales como, por ejemplo, amplificación, PCR, RT-PCR, protección de RNasa, transferencia Northern, espectrometría y otros métodos de hibridación.

Además, se divulga un ensayo de diagnóstico que puede usarse para detectar la sobreexpresión de polipéptidos de la proteína E en comparación con muestras normales de tejido de control para detectar la presencia de una infección, por ejemplo. Las técnicas de ensayo que pueden usarse para determinar los niveles de polipéptidos de la proteína E, en una muestra derivada de un huésped, tal como un material corporal, son bien conocidos por los expertos en la técnica. Tales métodos de ensayo incluyen radioinmunoensayos, ensayos de unión competitiva, análisis de transferencia Western, ensayos tipo sándwich de anticuerpos, detección de anticuerpos y ensayos ELISA.

Los polinucleótidos divulgados en este documento pueden usarse como componentes de matrices de polinucleótidos, preferiblemente matrices de alta densidad o rejillas. Estas matrices de alta densidad son particularmente útiles para propósitos de diagnóstico y pronóstico. Por ejemplo, un conjunto de puntos que comprenden cada uno un gen diferente, y que comprende además un polinucleótido o polinucleótidos de la invención, se puede usar para sondear, tal como usando hibridación o amplificación de ácidos nucleicos, usando sondas obtenidas o derivadas de una muestra corporal, para determinar la presencia de una secuencia de polinucleótidos particular o secuencia relacionada en un individuo. Tal presencia puede indicar la presencia de un patógeno, particularmente de *H. influenzae* no tipificable, y puede ser útil en el diagnóstico y/o pronóstico de enfermedad o un curso de la enfermedad. Se prefiere una rejilla que comprende un número de variantes de cualquier secuencia de polinucleótidos de la SEQ ID NO: 11. También se prefiere un número de variantes de una secuencia de polinucleótidos que codifican cualquier secuencia de polipéptido de la SEQ ID NO: 1-2 y 4-10.

Anticuerpos

Los polipéptidos y polinucleótidos de la divulgación o variantes de los mismos, o las células que expresan los mismos pueden usarse como inmunógenos para producir anticuerpos inmuno-específicos para tales polipéptidos o polinucleótidos respectivamente. Alternativamente, también pueden utilizarse mimótopos, particularmente mimótopos peptídicos, de epítomos dentro de la secuencia de polipéptidos como inmunógenos para producir anticuerpos inmuno-específicos para el polipéptido de la divulgación. El término "inmuno-específico" significa que los anticuerpos tienen una afinidad sustancialmente mayor por los polipéptidos de la divulgación que su afinidad por otros polipéptidos relacionados en la técnica anterior.

Los anticuerpos generados contra los polipéptidos o polinucleótidos de la divulgación pueden obtenerse administrando los polipéptidos y/o polinucleótidos de la invención, o fragmentos portadores de epítomos de uno o ambos, análogos de uno o ambos, o células que expresan uno o ambos, a un animal, preferiblemente no humano, utilizando protocolos de rutina. Para la preparación de anticuerpos monoclonales, se puede usar cualquier técnica conocida en el arte que proporcione anticuerpos producidos por cultivos continuos de líneas celulares. Los ejemplos incluyen diversas técnicas, tales como las de Kohler, G. y Milstein, C., Nature 256: 495-497 (1975); Kozbor et al., Immunology Today 4: 72 (1983); Cole et al., páginas 77-96 en MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc. (1985).

Las técnicas para la producción de anticuerpos de cadena sencilla (Patente de Estados Unidos No. 4.946.778) se pueden adaptar para producir anticuerpos de cadena sencilla para polipéptidos o polinucleótidos de esta divulgación. Además, pueden usarse ratones transgénicos u otros organismos o animales, tales como otros mamíferos, para expresar anticuerpos humanizados inmunoespecíficos para los polipéptidos o polinucleótidos de la invención.

Alternativamente, la tecnología de presentación en fagos puede utilizarse para seleccionar genes de anticuerpos con actividades de unión hacia un polipéptido de la invención, ya sea a partir de repertorios de genes v amplificados por PCR de linfocitos de seres humanos seleccionados por poseer bibliotecas de anti-proteína E o no modificadas (McCafferty, et al., (1990), Nature 348, 552-554; Marks, et al., (1992) Biotechnology 10, 779-783). La afinidad de estos anticuerpos también puede mejorarse mediante, por ejemplo, el arrastre de cadenas (Clackson et al., (1991) Nature 352: 628).

Los anticuerpos descritos anteriormente pueden emplearse para aislar o identificar clones que expresan los polipéptidos o polinucleótidos de la divulgación para purificar los polipéptidos o polinucleótidos mediante, por ejemplo, cromatografía de afinidad.

Por lo tanto, entre otros, los anticuerpos contra los polipéptidos de la proteína E o los polinucleótidos de la proteína E pueden emplearse para tratar infecciones, particularmente infecciones bacterianas.

Las variantes de polipéptidos incluyen variantes antigénicas, epitópicas o inmunológicamente equivalentes que forman un aspecto particular de esta divulgación.

Preferiblemente, el anticuerpo o variante del mismo se modifica para hacerlo menos inmunogénico en el individuo. Por ejemplo, si el individuo es humano, el anticuerpo puede preferiblemente "humanizarse", cuando la región o regiones determinantes de complementariedad del anticuerpo derivado de hibridoma se ha trasplantado en un anticuerpo monoclonal humano, por ejemplo, como se describe en Jones et al., (1986), Nature 321, 522-525 o Tempest et al., (1991) Biotechnology 9, 266-273.

Antagonistas y agonistas - Ensayos y moléculas

Los polipéptidos y polinucleótidos de la divulgación también se pueden usar para evaluar la unión de sustratos y ligandos de moléculas pequeñas en, por ejemplo, células, preparaciones libres de células, bibliotecas químicas y mezclas de productos naturales. Estos sustratos y ligandos pueden ser sustratos y ligandos naturales o pueden ser miméticos estructurales o funcionales. Véase, por ejemplo, Coligan et al., Current Protocols in Immunology 1 (2): capítulo 5 (1991).

Los métodos de detección pueden medir simplemente la unión de un compuesto candidato al polipéptido o polinucleótido, o a células o membranas que portan el polipéptido o polinucleótido, o una proteína de fusión del polipéptido por medio de un marcador asociado directa o indirectamente con el compuesto candidato. Alternativamente, el método de selección puede involucrar la competencia con un competidor etiquetado. Además, estos métodos de detección pueden evaluar si el compuesto candidato produce una señal generada por la activación o inhibición del polipéptido o polinucleótido, utilizando sistemas de detección apropiados para las células que comprenden el polipéptido o polinucleótido. Los inhibidores de activación generalmente se analizan en presencia de un agonista conocido y se observa el efecto sobre la activación por el agonista por la presencia del compuesto candidato. El polipéptido constitutivamente activo y/o los polipéptidos y polinucleótidos expresados constitutivamente pueden emplearse en métodos de detección de agonistas o inhibidores inversos, en ausencia de un agonista o inhibidor, al probar si el compuesto candidato da como resultado la inhibición de la activación del polipéptido o polinucleótido, como puede ser el caso. Además, los métodos de detección pueden comprender simplemente las etapas de mezclar un compuesto candidato con una solución que contiene un polipéptido o polinucleótido de la presente divulgación, para formar una mezcla, medir la actividad de los polipéptidos y/o polinucleótidos de la proteína E en la mezcla y comparar la actividad de los polipéptidos y/o polinucleótidos de la proteína E de la mezcla con un estándar. Las proteínas de fusión, tales como las hechas a partir de la porción Fc y los polipéptidos de la proteína E, como se describió anteriormente, también se pueden usar para ensayos de detección de alto rendimiento para identificar antagonistas del polipéptido de la presente divulgación, así como de filogenética y/o polipéptidos funcionalmente relacionados (véase D. Bennett et al., J Mol Recognition, 8: 52-58 (1995); y K. Johanson et al., J Biol Chem, 270 (16): 9459-9471 (1995)).

Los polinucleótidos, polipéptidos y anticuerpos que se unen y/o interactúan con un polipéptido de la presente divulgación también se pueden usar para configurar métodos de detección para detectar el efecto de compuestos añadidos en la producción de ARNm y/o polipéptidos en células. Por ejemplo, se puede construir un ensayo ELISA para medir niveles de polipéptidos secretados o asociados a células usando anticuerpos monoclonales y policlonales mediante métodos estándar conocidos en la técnica. Esto puede usarse para descubrir agentes que pueden inhibir o mejorar la producción de polipéptidos (también llamados antagonistas o agonistas, respectivamente) a partir de células o tejidos adecuadamente manipulados.

También se divulga un método de selección de compuestos para identificar aquellos que potencian (agonista) o bloquean (antagonista) la acción de polipéptidos o polinucleótidos de proteína E, particularmente aquellos compuestos que son bacteriostáticos y/o bactericidas. El método de detección puede involucrar técnicas de alto rendimiento. Por ejemplo, para detectar agonistas o antagonistas, una mezcla de reacción sintética, un compartimento celular, tal como una membrana, envoltura celular o pared celular, o una preparación de cualquiera de los mismos, que comprende polipéptidos de la proteína E y un sustrato o ligando marcado de dicho polipéptido se incuban en ausencia o en presencia de una molécula candidata que puede ser un agonista o antagonista de la proteína E. La capacidad de la molécula candidata para agonizar o antagonizar el polipéptido de la proteína E se refleja en la disminución de la unión del ligando marcado o en la disminución de la producción del producto a partir de dicho sustrato. Las moléculas que se unen gratuitamente, es decir, sin inducir los efectos del polipéptido de la proteína E, tienen más probabilidades de ser buenos antagonistas. Las moléculas que se unen bien y, según sea el caso, aumentan la velocidad de producción del producto a partir del sustrato, aumentan la transducción de señales o aumentan la actividad del canal químico son agonistas. La detección de la velocidad o nivel de, según sea el caso, la producción del producto a partir del sustrato, la transducción de señales o la actividad del canal químico puede mejorarse mediante el uso de un sistema informador. Los sistemas informadores que pueden ser útiles a este respecto incluyen, pero no se limitan a, colorimétrico sustrato marcado convertido en producto, un gen informador que responde a cambios en la actividad del polinucleótido o polipéptido de la proteína E y ensayos de unión conocidos en la técnica.

Otro ejemplo de un ensayo para agonistas de proteína E es un ensayo competitivo que combina proteína E y un agonista potencial con moléculas de unión a proteína E, moléculas de unión a proteína E recombinantes, sustratos o ligandos naturales, o miméticos de sustrato o ligando, en condiciones apropiadas para un ensayo de inhibición competitiva. La proteína E puede marcarse, tal como mediante radioactividad o un compuesto colorimétrico, de modo que el número de moléculas de proteína E unidas a una molécula de unión o convertidas en producto pueda determinarse con precisión para evaluar la efectividad del antagonista potencial.

Los antagonistas potenciales incluyen, entre otros, pequeñas moléculas orgánicas, péptidos, polipéptidos y anticuerpos que se unen a un polinucleótido y/o polipéptido de la invención y por lo tanto inhiben o extinguen su actividad o expresión. Los antagonistas potenciales también pueden ser moléculas orgánicas pequeñas, un péptido, un polipéptido tal como una proteína o anticuerpo estrechamente relacionado que se une a los mismos sitios en una molécula de unión, tal como una molécula de unión, sin inducir actividades inducidas por la proteína E, evitando así la acción o expresión de polipéptidos y/o polinucleótidos de proteína E al excluir los polipéptidos y/o polinucleótidos de proteína E de la unión.

Los antagonistas potenciales incluyen una molécula pequeña que se une y ocupa el sitio de unión del polipéptido evitando así la unión a moléculas de unión celular, de modo que se impide la actividad biológica normal. Los ejemplos de moléculas pequeñas incluyen, pero no se limitan a, moléculas orgánicas pequeñas, péptidos o moléculas similares a péptidos. Otros antagonistas potenciales incluyen moléculas antisentido (véase Okano, J. Neurochem. 56: 560 (1991); OLIGODEOXYNUCLEOTIDES AS ANTISENSE INHIBITORS OF GENE EXPRESSION, CRC Press, Boca Raton, FL (1988), para una descripción de estas moléculas). Los antagonistas potenciales preferidos incluyen compuestos relacionados y variantes de la proteína E.

En otro aspecto, la presente divulgación se refiere a proteínas de fusión solubles genéticamente modificadas que comprenden un polipéptido de la presente divulgación, o un fragmento del mismo, y varias porciones de las regiones constantes de cadenas pesadas o ligeras de inmunoglobulinas de diversas subclases (IgG, IgM, IgA, IgE). Se prefiere como inmunoglobulina la parte constante de la cadena pesada de la IgG humana, particularmente la IgG1, en la que la fusión tiene lugar en la región bisagra. En una realización particular, la parte Fc se puede eliminar simplemente mediante la incorporación de una secuencia de escisión que se puede escindir con el factor de coagulación sanguínea Xa. Además, esta divulgación se refiere a procesos para la preparación de estas proteínas de fusión mediante ingeniería genética, y al uso de las mismas para la detección, diagnóstico y terapia con fármacos. Un aspecto adicional de la divulgación también se refiere a polinucleótidos que codifican tales proteínas de fusión. Se pueden encontrar ejemplos de tecnología de proteínas de fusión en las solicitudes internacionales de patente Nos. WO94/29458 y WO94/22914.

Cada una de las secuencias de polinucleótidos proporcionadas en el presente documento pueden usarse en el descubrimiento y desarrollo de compuestos antibacterianos. La proteína codificada, después de la expresión, puede usarse como un objetivo para la detección de fármacos antibacterianos. Además, las secuencias polinucleotídicas que codifican las regiones amino terminales de la proteína codificada o Shine-Dalgarno u otras secuencias facilitadoras de la traducción del ARNm respectivo pueden usarse para construir secuencias antisentido para controlar la expresión de la secuencia codificante de interés.

La divulgación también proporciona el uso del polipéptido, polinucleótido, agonista o antagonista de la invención para interferir con la interacción física inicial entre un patógeno o patógenos y un huésped eucariota, preferiblemente mamífero, responsable de las secuelas de la infección. En particular, las moléculas de la divulgación pueden usarse: en la prevención de la adhesión de bacterias, en particular bacterias grampositivas y/o gramnegativas, a proteínas de matriz extracelular eucariotas, preferiblemente de mamífero, en dispositivos permanentes o a proteínas de matriz extracelular en heridas; para bloquear la adhesión bacteriana entre proteínas de matriz extracelular eucariotas, preferiblemente de mamíferos, y proteínas de proteína E bacterianas que median el daño tisular y/o; para bloquear la

progresión normal de la patogénesis en infecciones iniciadas que no sean por la implantación de dispositivos permanentes o por otras técnicas quirúrgicas.

5 De acuerdo con aún otro aspecto de la divulgación, se proporcionan agonistas y antagonistas de la proteína E, preferiblemente agonistas y antagonistas bacteriostáticos o bactericidas.

Los antagonistas y agonistas pueden emplearse, por ejemplo, para prevenir, inhibir y/o tratar enfermedades.

10 En un aspecto adicional, la presente divulgación se refiere a mimótopos del polipéptido de la divulgación. Un mimótopo es una secuencia peptídica, suficientemente similar al péptido nativo (secuencial o estructuralmente), que es capaz de ser reconocida por anticuerpos que reconocen el péptido nativo; o es capaz de generar anticuerpos que reconocen el péptido nativo cuando se acopla a un vehículo adecuado.

15 Los mimótopos de péptidos pueden diseñarse para un propósito particular mediante la adición, eliminación o sustitución de aminoácidos elegidos. Por lo tanto, los péptidos pueden modificarse con el fin de facilitar la conjugación con un vehículo proteico. Por ejemplo, puede ser deseable que algunos métodos de conjugación química incluyan una cisteína terminal. Además, puede ser deseable que los péptidos conjugados con un vehículo proteico incluyan un terminal hidrófobo distal del terminal conjugado del péptido, de modo que el extremo libre no conjugado del péptido permanezca asociado con la superficie de la proteína portadora. De este modo, presenta el péptido en una conformación que se parece más a la del péptido que se encuentra en el contexto de la molécula nativa completa. Por ejemplo, los péptidos pueden alterarse para tener una cisteína en el terminal N y una cola amidada hidrófoba en el terminal C. Alternativamente, la adición o sustitución de una forma D-estereoisómera de uno o más de los aminoácidos (secuencias inversas) se puede realizar para crear un derivado beneficioso, por ejemplo, para mejorar la estabilidad del péptido. Los mimótopos también pueden ser secuencias retro de las secuencias de péptidos naturales, ya que la orientación de la secuencia se invierte. Los mimótopos también pueden ser de carácter retroinverso. Los péptidos retro, inverso y retroinverso se describen en los documentos WO 95/24916 y WO 94/05311.

30 Alternativamente, los mimótopos peptídicos pueden identificarse usando anticuerpos que son capaces de unirse a los polipéptidos de la presente invención usando técnicas tales como la tecnología de presentación en fagos (documento EP 0 552 267 B1). Esta técnica genera una gran cantidad de secuencias peptídicas que imitan la estructura de los péptidos nativos y, por lo tanto, son capaces de unirse a anticuerpos antipéptido nativo, pero pueden no necesariamente compartir homología de secuencia significativa con el polipéptido nativo.

35 Vacunas

Otro aspecto de la divulgación se refiere a un método para inducir una respuesta inmunológica en un individuo, particularmente un mamífero, preferiblemente humanos, que comprende inocular al individuo con polinucleótido y/o polipéptido de proteína E, o un fragmento o variante del mismo, adecuado para producir una respuesta inmune de anticuerpos y/o células T para proteger a dicho individuo de la infección, particularmente infección bacteriana y más particularmente infección por *H. influenzae* no tipificable. También se proporcionan métodos mediante los cuales dicha respuesta inmunológica ralentiza la replicación bacteriana. Otro aspecto más de la divulgación se refiere a un método para inducir una respuesta inmunológica en un individuo que comprende administrar a dicho individuo un vector de ácido nucleico, secuencia o ribozima para dirigir la expresión de polinucleótidos y/o polipéptidos de proteína E, o un fragmento o una variante de los mismos, para expresar polinucleótidos y/o polipéptidos de proteína E, o un fragmento o una variante de los mismos *in vivo* para inducir una respuesta inmunológica, tal como, producir respuesta inmune de anticuerpos y/o células T, que incluye, por ejemplo, células T productoras de citocinas o células T citotóxicas, para proteger a dicho individuo, preferiblemente un ser humano, de la enfermedad, ya sea que la enfermedad ya esté establecida dentro del individuo o no. Un ejemplo de administración del gen es acelerándolo en las células deseadas como un recubrimiento sobre partículas o de otro modo. Dicho vector de ácido nucleico puede comprender ADN, ARN, una ribozima, un ácido nucleico modificado, un híbrido de ADN/ARN, un complejo de ADN-proteína o un complejo de ARN-proteína.

50 Un aspecto adicional de la divulgación se refiere a una composición inmunológica que cuando se introduce en un individuo, preferiblemente un ser humano, capaz de haber inducido dentro de él una respuesta inmunológica, induce una respuesta inmunológica en dicho individuo a un polinucleótido y/o polipéptido de proteína E codificado a partir del mismo, en el que la composición comprende un polinucleótido y/o polipéptido de proteína E recombinante codificado a partir del mismo y/o comprende ADN y/o ARN que codifica y expresa un antígeno de dicho polinucleótido y/o polipéptido de proteína E, codificado a partir del mismo u otro polipéptido de la divulgación. La respuesta inmunológica puede usarse terapéutica o profilácticamente y puede tomar la forma de inmunidad de anticuerpo y/o inmunidad celular, tal como inmunidad celular que surge de células T CTL o CD4+.

60 Los polipéptidos de la proteína E o un fragmento de los mismos pueden fusionarse con coproteína o una fracción química que puede o no producir anticuerpos por sí misma, pero que es capaz de estabilizar la primera proteína y producir una proteína fusionada o modificada que tendrá propiedades antigénicas y/o inmunogénicas, y preferiblemente propiedades protectoras. Por lo tanto, la proteína recombinante fusionada, preferiblemente comprende además una coproteína antigénica, tal como lipoproteína o proteína D de *Haemophilus influenzae* (EP 594610), glutatión-S-transferasa (GST) o beta-galactosidasa, o cualquier otra coproteína relativamente grande que solubiliza la

proteína y facilita su producción y purificación. Además, la coproteína puede actuar como un adyuvante en el sentido de proporcionar una estimulación generalizada del sistema inmune del organismo que recibe la proteína. La coproteína puede estar unida al terminal amino o carboxilo de la primera proteína.

- 5 En una composición de la vacuna de acuerdo con la divulgación, un polipéptido y/o polinucleótido de proteína E, o un fragmento, o un mimótopo, o una variante del mismo pueden estar presentes en un vector, tal como los vectores recombinantes vivos descritos anteriormente, por ejemplo, vectores bacterianos vivos

10 También son adecuados los vectores no vivos para los polipéptidos de la proteína E, por ejemplo, vesículas o "ampollas" de la membrana externa bacteriana. Las ampollas de la OM se derivan de la membrana externa de la membrana de dos capas de bacterias Gram negativas y se han documentado en muchas bacterias Gram negativas (Zhou, L et al., 1998. FEMS Microbiol. Lett. 163: 223-228) incluyendo *C. trachomatis* y *C. psittaci*. Una lista no exhaustiva de patógenos bacterianos que se reportan que producen ampollas también incluye: *Bordetella pertussis*, *Borrelia burgdorferi*, *Brucella melitensis*, *Brucella ovis*, *Escherichia coli*, *Haemophilus influenzae*, *Legionella pneumophila*, *Moraxella catarrhalis*, *Neisseria gonorrhoeae*, *Neisseria meningitidis*, *Pseudomonas aeruginosa* y *Yersinia enterocolitica*.

20 Las ampollas tienen la ventaja de proporcionar proteínas de membrana externa en su conformación nativa y, por lo tanto, son particularmente útiles para las vacunas. Las ampollas también se pueden mejorar para el uso de vacunas mediante la modificación de la bacteria para modificar la expresión de una o más moléculas en la membrana externa. Por lo tanto, por ejemplo, la expresión de una proteína inmunogénica deseada en la membrana externa, tal como los polipéptidos de la proteína E, puede introducirse o sobreregularse (por ejemplo, alterando el promotor). En cambio, o además, la expresión de moléculas de membrana externa que no son relevantes (por ejemplo, antígenos no protectores o inmunodominantes pero proteínas variables) o perjudiciales (por ejemplo, moléculas tóxicas tales como LPS o posibles inductores de una respuesta autoinmune) pueden ser subreguladas. Estos enfoques se analizan con más detalle a continuación.

30 Las regiones flanqueantes no codificantes de los genes de la proteína E contienen elementos reguladores importantes en la expresión del gen. Esta regulación tiene lugar tanto a nivel transcripcional como de traducción. La secuencia de estas regiones, secuencia arriba o secuencia abajo del marco de lectura abierto del gen, puede obtenerse mediante secuenciación de ADN. Esta información de secuencia permite la determinación de motivos reguladores potenciales, tal como los diferentes elementos promotores, secuencias terminadoras, elementos de secuencia inducibles, represores, elementos responsables de la variación de fase, la secuencia de Shine-Dalgarno, regiones con una estructura secundaria potencial involucrada en la regulación, así como otros tipos de motivos o secuencias reguladoras. Esta secuencia es un aspecto adicional de la divulgación.

40 Esta información de secuencia permite la modulación de la expresión natural de los genes de la proteína E. La sobreregularización de la expresión génica se puede lograr alterando el promotor, la secuencia de Shine-Dalgarno, los elementos represores u operadores potenciales, o cualquier otro elemento involucrado. Del mismo modo, la subregulación de la expresión se puede lograr mediante tipos similares de modificación. Alternativamente, al cambiar las secuencias de variación de fase, la expresión del gen se puede poner bajo control de variación de fase, o se puede desacoplar de esta regulación. En otro enfoque, la expresión del gen se puede poner bajo el control de uno o más elementos inducibles que permiten la expresión regulada. Los ejemplos de dicha regulación incluyen, entre otros, la inducción por cambio de temperatura, la adición de sustratos inductores como carbohidratos seleccionados o sus derivados, oligoelementos, vitaminas, cofactores, iones metálicos, etc.

50 Las modificaciones descritas anteriormente se pueden introducir por varios medios diferentes. La modificación de las secuencias implicadas en la expresión génica puede llevarse a cabo *in vivo* mediante mutagénesis aleatoria seguida de selección para el fenotipo deseado. Otro enfoque consiste en aislar la región de interés y modificarla mediante mutagénesis aleatoria, o mutagénesis de reemplazo, inserción o eliminación dirigida al sitio. La región modificada se puede reintroducir en el genoma bacteriano mediante recombinación homóloga, y se puede evaluar el efecto sobre la expresión génica. En otro enfoque, el conocimiento de la secuencia de la región de interés puede usarse para reemplazar o eliminar todo o parte de las secuencias reguladoras naturales. En este caso, la región reguladora objetivo está aislada y modificada para contener los elementos reguladores de otro gen, una combinación de elementos reguladores de diferentes genes, una región reguladora sintética o cualquier otra región reguladora, o para eliminar partes seleccionadas de secuencias reguladoras de tipo silvestre. Estas secuencias modificadas pueden reintroducirse en la bacteria mediante recombinación homóloga en el genoma. Una lista no exhaustiva de promotores preferidos que podrían usarse para la sobreregularización de la expresión génica incluye los promotores *porA*, *porB*, *lbpB*, *tbpB*, *p110*, *lst*, *hpuAB* de *N. meningitidis* o *N. gonorrhoeae*; *ompCD*, *copB*, *lbpB*, *ompE*, *UspA1*; *UspA2*; *TbpB* de *M. Catarrhalis*; *p1*, *p2*, *p4*, *p5*, *p6*, *lpD*, *tbpB*, *D15*, *Hia*, *Hmw1*, *Hmw2* de *H. influenzae*.

60 En un ejemplo, la expresión del gen puede modularse intercambiando su promotor con un promotor más fuerte (mediante el aislamiento de la secuencia en sentido 5' del gen, la modificación *in vitro* de esta secuencia y la reintroducción en el genoma mediante recombinación homóloga). La expresión sobreregularada se puede obtener tanto en la bacteria como en las vesículas de la membrana externa desprendidas (o elaboradas) de la bacteria.

En otros ejemplos, los enfoques descritos pueden usarse para generar cepas bacterianas recombinantes con características mejoradas para aplicaciones de vacunas. Estas pueden ser, entre otras, cepas atenuadas, cepas con expresión incrementada de antígenos seleccionados, cepas con inactivaciones (o expresión disminuida) de genes que interfieren con la respuesta inmune, cepas con expresión modulada de proteínas inmunodominantes, cepas con desprendimiento modulado de vesículas de membrana externa.

Por lo tanto, también es proporcionada por la divulgación una región secuencia arriba modificada de los genes de la proteína E, cuya región secuencia arriba modificada contiene un elemento regulador heterólogo que altera el nivel de expresión de las proteínas de la proteína E ubicadas en la membrana externa. La región secuencia arriba de acuerdo con este aspecto de la divulgación incluye la secuencia en dirección 5' de los genes de la proteína E. La región secuencia arriba comienza inmediatamente secuencia arriba de los genes de la proteína E y continúa generalmente hasta una posición no mayor a aproximadamente 1000 pb secuencia arriba del gen del codón de inicio ATG. En el caso de un gen ubicado en una secuencia policistrónica (operón), la región secuencia arriba puede comenzar inmediatamente antes del gen de interés, o antes del primer gen en la región secuencia arriba puede comenzar inmediatamente antes del gen de interés, o antes del primer gen en el operón. Preferiblemente, una región secuencia arriba modificada de acuerdo con este aspecto de la divulgación contiene un promotor heterólogo en una posición entre 500 y 700 pb secuencia arriba del ATG.

El uso de las regiones secuencia arriba descritas para sobrerregular la expresión de los genes de la proteína E, un proceso para lograr esto a través de la recombinación homóloga (por ejemplo, como se describe en el documento WO 01/09350, un vector que comprende la secuencia en dirección 5' adecuada para este propósito, y también se divulga una célula huésped así alterada.

Por lo tanto, la divulgación proporciona polipéptidos de la proteína E, en una ampolla bacteriana modificada. La divulgación proporciona además células huésped modificadas capaces de producir los vectores de ampolla basados en la membrana no viva. La divulgación proporciona además vectores de ácido nucleico que comprenden los genes de la proteína E que tienen una región secuencia arriba modificada que contiene un elemento regulador heterólogo.

Se proporcionan además procesos para preparar las células huésped y ampollas bacterianas de acuerdo con la divulgación.

También se proporcionan mediante esta divulgación composiciones, particularmente composiciones de las vacunas, y métodos que comprenden los polipéptidos y/o polinucleótidos de la divulgación y secuencias de ADN inmunoestimulantes, tales como las descritas en Sato, Y. et al., Science 273: 352 (1996).

Además, en esta divulgación se proporcionan métodos que usan el polinucleótido descrito o fragmentos particulares del mismo, que se ha demostrado que codifican regiones no variables de proteínas de la superficie celular bacteriana, en constructos de polinucleótidos usados en tales experimentos de inmunización genética en modelos animales de infección con *H. influenzae* no tipificable. Tales experimentos serán particularmente útiles para identificar epítopos de proteínas capaces de provocar una respuesta inmune profiláctica o terapéutica. Se cree que este enfoque permitirá la posterior preparación de anticuerpos monoclonales de particular valor, derivados del órgano requerido del animal que resiste o elimina con éxito la infección, para el desarrollo de agentes profilácticos o tratamientos terapéuticos de infección bacteriana, particularmente infección por *H. influenzae* no tipificable en mamíferos, particularmente en humanos.

La divulgación también incluye una formulación de vacuna que comprende un polipéptido y/o polinucleótido recombinante inmunogénico de la divulgación junto con un vehículo adecuado, tal como un vehículo farmacéuticamente aceptable. Dado que los polipéptidos y polinucleótidos pueden descomponerse en el estómago, cada uno se administra preferiblemente por vía parenteral, incluida, por ejemplo, la administración que es subcutánea, intramuscular, intravenosa o intradérmica. Las formulaciones adecuadas para administración parenteral incluyen soluciones de inyección estériles acuosas y no acuosas que pueden contener antioxidantes, tampones, compuestos bacteriostáticos y solutos que hacen que la formulación sea isotónica con el fluido corporal, preferiblemente la sangre, del individuo; y suspensiones estériles acuosas y no acuosas que pueden incluir agentes de suspensión o agentes espesantes. Las formulaciones pueden presentarse en envases de dosis unitarias o multidosis, por ejemplo, ampollas selladas y viales, y pueden almacenarse en una condición liofilizada que requiera solo la adición del vehículo líquido estéril inmediatamente antes de su uso.

La formulación de la vacuna también puede incluir sistemas adyuvantes para mejorar la inmunogenicidad de la formulación. Preferiblemente, el sistema adyuvante genera preferiblemente un tipo de respuesta TH1.

Una respuesta inmune puede distinguirse ampliamente en dos categorías extremas, siendo una respuesta inmune humoral o mediada por células (caracterizada tradicionalmente por mecanismos de protección de anticuerpos y efectores celulares respectivamente).

Las respuestas inmunes extremas de tipo TH1 pueden caracterizarse por la generación de linfocitos T citotóxicos restringidos por haplotipo específicos del antígeno y respuestas de células asesinas naturales. En los ratones, las

respuestas de tipo TH1 a menudo se caracterizan por la generación de anticuerpos del subtipo IgG2a, mientras que en humanos corresponden a anticuerpos de tipo IgG1. Las respuestas inmunes de tipo TH2 se caracterizan por la generación de una amplia gama de isotipos de inmunoglobulina, incluidos en IgG1, IgA e IgM de ratón.

- 5 Se puede considerar que la fuerza impulsora detrás del desarrollo de estos dos tipos de respuestas inmunes son las citocinas. Los altos niveles de citocinas de tipo TH1 tienden a favorecer la inducción de respuestas inmunes mediadas por células al antígeno dado, mientras que los altos niveles de citocinas de tipo TH2 tienden a favorecer la inducción de respuestas inmunes humorales al antígeno.
- 10 La distinción de las respuestas inmunes de tipo TH1 y TH2 no es absoluta. En realidad, un individuo apoyará una respuesta inmune que se describe como predominantemente TH1 o predominantemente TH2. Sin embargo, a menudo es conveniente considerar las familias de citocinas en términos de lo descrito en clones de células T CD4+ve murinas por Mosmann y Coffman (Mosmann, TR y Coffman, RL (1989) TH1 and TH2 cells: different patterns of lymphokine secretion lead to different functional properties. Annual Review of Immunology, 7, páginas 145-173). Tradicionalmente,
- 15 las respuestas de tipo TH1 están asociadas con la producción de las citocinas INF- γ e IL-2 por los linfocitos T. Otras células citocinas a menudo directamente asociadas con la inducción de respuestas inmunes de tipo TH1 no son producidas por las células T, tales como IL-12. En contraste, las respuestas de tipo TH2 están asociadas con la secreción de IL-4, IL-5, IL-6 e IL-13.
- 20 Se sabe que ciertos adyuvantes de vacuna son particularmente adecuados para la estimulación de respuestas de citocinas de tipo TH1 o TH2. Tradicionalmente, los mejores indicadores del equilibrio TH1:TH2 de la respuesta inmune después de una vacunación o infección incluyen la medición directa de la producción de citocinas TH1 o TH2 por los linfocitos T *in vitro* después de la reestimulación con antígeno, y/o la medición de la relación IgG1:IgG2a de respuestas de anticuerpos específicos de antígeno.
- 25 Por lo tanto, un adyuvante de tipo TH1 es aquel que estimula preferiblemente poblaciones de células T aisladas para producir altos niveles de citocinas de tipo TH1 cuando se reestiman con antígeno *in vitro*, y promueve el desarrollo tanto de linfocitos T citotóxicos CD8+ como de respuestas de inmunoglobulina específicas de antígeno asociadas con el isotipo tipo TH1.
- 30 Los adyuvantes que son capaces de estimulación preferencial de la respuesta celular TH1 se describen en la solicitudes internacionales de patente Nos. WO 94/00153 y WO 95/17209.
- 35 El lípido A monofosforil 3-des-O-acilado (3D-MPL) es uno de tales adyuvantes. Este se conoce a partir del documento GB 2220211 (Ribi). Químicamente es una mezcla del lípido A monofosforil 3-des-O-acilado con 4, 5 o 6 cadenas aciladas y es fabricado por Ribi Immunochem, Montana. Una forma preferida de lípido A monofosforil 3-des-O-acilado se divulga en la patente europea 0 689 454 B1 (SmithKline Beecham Biologicals SA).
- 40 Preferiblemente, las partículas de 3D-MPL son lo suficientemente pequeñas como para ser esterilizadas a través de una membrana de 0.22 micras (patente europea No. 0 689 454).
- 45 3D-MPL estará presente en el intervalo de 10 pg - 100 μ g preferiblemente 25 - 50 pg por dosis en el que el antígeno típicamente estará presente en un intervalo de 2 - 50 pg por dosis.
- Otro adyuvante preferido comprende QS21, una fracción no tóxica purificada por HPLC derivada de la corteza de Quillaja Saponaria Molina. Opcionalmente, esto puede mezclarse con lípido A monofosforil 3-des-O-acilado (3D-MPL), opcionalmente junto con un vehículo.
- 50 El método de producción de QS21 se divulga en la patente de Estados Unidos No. 5.057.540.
- 55 Se han descrito previamente formulaciones adyuvantes no reactogénicas que contienen QS21 (documento WO 96/33739). Se ha demostrado que dichas formulaciones que comprenden QS21 y colesterol son adyuvantes exitosos estimulantes de TH1 cuando se formulan junto con un antígeno.
- 60 Los adyuvantes adicionales que son estimuladores preferenciales de la respuesta celular TH1 incluyen oligonucleótidos inmunomoduladores, por ejemplo secuencias CpG no metiladas como se divulga en el documento WO 96/02555.
- También se contempla que las combinaciones de diferentes adyuvantes estimulantes de TH1, tales como las mencionadas anteriormente, proporcionan un adyuvante que es un estimulador preferencial de la respuesta celular TH1. Por ejemplo, QS21 puede formularse junto con 3D-MPL. La relación de QS21:3D-MPL será típicamente del orden de 1:10 a 10:1; preferiblemente 1:5 a 5:1 y a menudo sustancialmente 1:1. El intervalo preferido para una sinergia óptima es 2,5:1 a 1:1 de 3D-MPL:QS21.

Preferiblemente, también está presente un vehículo en la composición de la vacuna de acuerdo con la divulgación. El vehículo puede ser una emulsión de aceite en agua, o una sal de aluminio, tal como fosfato de aluminio o hidróxido de aluminio.

5 Una emulsión preferida de aceite en agua comprende un aceite metabolizable, tal como escualeno, alfa tocoferol y Tween 80. En un aspecto particularmente preferido, los antígenos en la composición de la vacuna de acuerdo con la divulgación se combinan con QS21 y 3D-MPL en dicha emulsión. Además, la emulsión de aceite en agua puede contener Span 85 y/o lecitina y/o tricaprilina.

10 Típicamente para administración humana, QS21 y 3D-MPL estarán presentes en una vacuna en el intervalo de 1 µg - 200 pg, tal como 10 - 100 pg, preferiblemente 10 pg - 50 pg por dosis. Típicamente, el aceite en agua comprenderá del 2 al 10% de escualeno, del 2 al 10% de alfa tocoferol y del 0,3 al 3% de Tween 80. Preferiblemente, la proporción de escualeno:alfa tocoferol es igual o inferior a 1, ya que esto proporciona una emulsión más estable. Span 85 también puede estar presente en un nivel del 1%. En algunos casos puede ser ventajoso que las vacunas de la presente divulgación contengan además un estabilizador.

15 Las emulsiones de aceite en agua no tóxicas contienen preferiblemente un aceite no tóxico, por ejemplo, escualano o escualeno, un emulsionante, por ejemplo, Tween 80, en un vehículo acuoso. El vehículo acuoso puede ser, por ejemplo, solución salina tamponada con fosfato.

20 Una formulación adyuvante particularmente potente que implica QS21, 3D-MPL y tocoferol en una emulsión de aceite en agua se describe en el documento WO 95/17210. Los fragmentos/péptidos preferidos se describen en el Ejemplo 10.

25 La presente divulgación también proporciona una composición de la vacuna polivalente que comprende una formulación de vacuna de la invención en combinación con otros antígenos, en particular antígenos útiles para tratar la otitis media. Tal composición de la vacuna polivalente puede incluir un adyuvante inductor de TH-1 como se describió anteriormente en este documento.

30 En una realización preferida, los polipéptidos, fragmentos e inmunógenos de la invención se formulan con uno o más de los siguientes grupos de antígenos: a) uno o más polisacáridos capsulares neumocócicos (ya sea simples o conjugados con una proteína transportadora); b) uno o más antígenos que pueden proteger a un huésped contra la infección por *M. catarrhalis*; c) uno o más antígenos proteicos que pueden proteger a un huésped contra la infección por *Streptococcus pneumoniae*; d) uno o más antígenos proteicos de *Haemophilus influenzae* no tipificables adicionales; e) uno o más antígenos que pueden proteger a un huésped contra el VSR; y f) uno o más antígenos que pueden proteger a un huésped contra el virus de la influenza. Combinaciones con: grupos a) y b); b) y c); b), d), y a) y/o c); se prefieren b), d), e), f) y a) y/o c). Dichas vacunas pueden usarse ventajosamente como vacunas globales contra la otitis media.

35 40 Los antígenos polisacáridos capsulares neumocócicos se seleccionan preferiblemente de los serotipos 1, 2, 3, 4, 5, 6B, 7F, 8, 9N, 9V, 10A, 11A, 12F, 14, 15B, 17F, 18C, 19A, 19F, 20, 22F, 23F y 33F (más preferiblemente de los serotipos 1, 3, 4, 5, 6B, 7F, 9V, 14, 18C, 19F y 23F).

45 Los antígenos de proteína neumocócica preferidos son aquellas proteínas neumocócicas que están expuestas en la superficie externa del neumococo (capaces de ser reconocidas por el sistema inmune del huésped durante al menos parte del ciclo de vida del neumococo), o son proteínas que se secretan o liberan por el neumococo. Lo más preferiblemente, la proteína es una toxina, adhesina, transductor de señal de 2 componentes o lipoproteína de *Streptococcus pneumoniae*, o fragmentos de la misma. Las proteínas particularmente preferidas incluyen, pero no se limitan a: neumolisina (preferiblemente desintoxicada por tratamiento químico o mutación) [Mitchell et al., Nucleic Acids Res. 11 de julio de 1990; 18 (13): 4010 "Comparison of pneumolysin genes and proteins from *Streptococcus pneumoniae* types 1 and 2", Mitchell et al., Biochim Biophys Acta. 23 de enero de 1989; 1007(1): 67-72 "Expression of the pneumolysin gene in *Escherichia coli*: rapid purification and biological properties", WO 96/05859 (A. Cyanamid), WO 90/06951 (Paton et al), WO 99/03884 (NAVA)]; PspA y variantes de eliminación transmembrana de los mismos (WO 92/14488; WO 99/53940; US 5804193 - Briles et al.); PspC y variantes de eliminación transmembrana de los mismos (WO 99/53940; WO 97/09994 - Briles et al); PsaA y sus variantes de eliminación transmembrana (Berry & Paton, Infect Immun, diciembre de 1996; 64 (12): 5255-62 "Sequence heterogeneity of PsaA, a 37-kilodalton putative adhesin essential for virulence of *Streptococcus pneumoniae*"); proteínas de unión a colina neumocócicas y variantes de eliminación transmembrana de las mismas; CbpA y variantes de eliminación transmembrana de las mismas (WO 97/41151; WO 99/51266); Gliceraldehído-3-fosfato-deshidrogenasa (Infect. Immun. 1996 64: 3544); HSP70 (WO 96/40928); PcpA (Sanchez-Beato et al. FEMS Microbiol Lett 1998, 164: 207-14); Proteína tipo M, solicitud de patente SB No. EP 0837130; y adhesina 18627 (solicitud de patente SB No. EP 0834568). Otros antígenos proteicos neumocócicos preferidos son los divulgados en el documento WO 98/18931, particularmente los seleccionados en los documentos WO 98/18930 y PCT/US99/30390.

65 Los antígenos de proteína preferidos de *Moraxella catarrhalis* que se pueden incluir en una vacuna combinada (especialmente para la prevención de otitis media) son: OMP106 [WO 97/41731 (Antex) y WO 96/34960 (PMC)];

OMP21; LbpA y/o LbpB [WO 98/55606 (PMC)]; TbpA y/o TbpB [WO 97/13785 y WO 97/32980 (PMC)]; CopB [Helminen ME, et al. (1993) Infectar. Immun 61: 2003-2010]; UspA1 y/o UspA2 [WO 93/03761 (Universidad de Texas)]; OmpCD; HasR (PCT/EP99/03824); PilQ (PCT/EP99/03823); OMP85 (PCT/EP00/01468); lipo06 (GB 9917977.2); lipo10 (GB 9918208.1); lipo11 (GB 9918302.2); lipo18 (GB 9918038.2); P6 (PCT/EP99/03038); D15 (PCT/EP99/03822); OmpIA1 (PCT/EP99/06781); Hly3 (PCT/EP99/03257); y OmpE.

Los antígenos de proteína de *Haemophilus influenzae* no tipificables preferidos adicionales que pueden incluirse en una vacuna combinada (especialmente para la prevención de otitis media) incluyen: proteína Fimbrina [(US 5766608 - Ohio State Research Foundation)] y fusiones que comprenden péptidos de la misma [por ejemplo, fusiones de péptidos LB1 (f); US 5843464 (OSU) o WO 99/64067]; OMP26 [WO 97/01638 (Cortecs)]; P6 [EP 281673 (Universidad Estatal de Nueva York)]; proteína D (EP 594610); TbpA y/o TbpB; Hia Hsf; Hin47; Hif Hmw1; Hmw2; Hmw3; Hmw4; Hap; D15 (WO 94/12641); P2; P5 (WO 94/26304); NlpC2 (BASB205) [WO 02/30971]; Slp (BASB203) [WO 02/30960]; e iOMP1681 (BASB210) [WO 02/34772].

Los antígenos de virus de la influenza preferidos incluyen virus enteros, vivos o inactivados, virus de la influenza divididos, cultivados en huevos o células MDCK, o células Vero o virosomas completos de la gripe (como lo describe R. Gluck, Vaccine, 1992, 10, 915-920) o proteínas purificadas o recombinantes de las mismas, tales como proteínas HA, NP, NA o M, o combinaciones de las mismas.

Los antígenos de RSV (virus sincitial respiratorio) preferidos incluyen la glicoproteína F, la glicoproteína G, la proteína HN o sus derivados.

Composiciones, kits y administración

En un aspecto adicional de la divulgación, se proporcionan composiciones que comprenden polinucleótidos de proteína E y/o polipéptidos de proteína E para administración a una célula o a un organismo multicelular.

La divulgación también se refiere a composiciones que comprenden un polinucleótido y/o un polipéptido discutidos en el presente documento o sus agonistas o antagonistas. Los polipéptidos y polinucleótidos de la divulgación pueden emplearse en combinación con un vehículo o vehículos no estériles o estériles para usar con células, tejidos u organismos, tales como un vehículo farmacéutico adecuado para la administración a un individuo. Dichas composiciones comprenden, por ejemplo, un aditivo para el medio o una cantidad terapéuticamente efectiva de un polipéptido y/o polinucleótido de la divulgación y un vehículo o excipiente farmacéuticamente aceptable. Dichos vehículos pueden incluir, pero no se limitan a, solución salina, solución salina tamponada, dextrosa, agua, glicerol, etanol y combinaciones de los mismos. La formulación debe adaptarse al modo de administración. La invención se refiere además a paquetes y kits de diagnóstico y farmacéuticos que comprenden uno o más recipientes llenos con uno o más de los ingredientes de las composiciones mencionadas anteriormente.

Los polipéptidos, polinucleótidos y otros compuestos pueden emplearse solos o en combinación con otros compuestos, tales como compuestos terapéuticos.

Las composiciones farmacéuticas pueden administrarse de cualquier manera eficaz y conveniente que incluye, por ejemplo, la administración por vía tópica, oral, anal, vaginal, intravenosa, intraperitoneal, intramuscular, subcutánea, intranasal o intradérmica, entre otras.

En terapia o como profiláctico, el agente activo puede administrarse a un individuo como una composición inyectable, por ejemplo como una dispersión acuosa estéril, preferiblemente isotónica.

En otro aspecto, la presente divulgación proporciona composiciones farmacéuticas que comprenden una cantidad terapéuticamente efectiva de un polipéptido y/o polinucleótido, tal como la forma soluble de un polipéptido y/o polinucleótido de la presente divulgación, péptido agonista o antagonista o compuesto de molécula pequeña, en combinación con un vehículo o excipiente farmacéuticamente aceptable. Dichos vehículos incluyen, pero no se limitan a, solución salina, solución salina tamponada, dextrosa, agua, glicerol, etanol y combinaciones de los mismos. La divulgación se refiere además a paquetes y kits farmacéuticos que comprenden uno o más recipientes llenos con uno o más de los ingredientes de las composiciones mencionadas anteriormente. Los polipéptidos, polinucleótidos y otros compuestos de la presente divulgación pueden emplearse solos o en combinación con otros compuestos, tales como compuestos terapéuticos.

La composición se adaptará a la ruta de administración, por ejemplo, por una ruta sistémica u oral. Las formas preferidas de administración sistémica incluyen inyección, típicamente por inyección intravenosa. Se pueden usar otras vías de inyección, tales como la subcutánea, intramuscular o intraperitoneal. Los medios alternativos para la administración sistémica incluyen la administración transmucosa y transdérmica usando penetrantes tales como sales biliares o ácidos fusídicos u otros detergentes. Además, si un polipéptido u otros compuestos de la presente divulgación se pueden formular en una formulación entérica o encapsulada, también puede ser posible la administración oral. La administración de estos compuestos también puede ser tópica y/o localizada, en forma de ungüentos, pastas, geles, soluciones, polvos y similares.

5 Para la administración a mamíferos, y particularmente a humanos, se espera que el nivel de dosificación diaria del agente activo sea de 0,01 mg/kg a 10 mg/kg, típicamente alrededor de 1 mg/kg. En cualquier caso, el médico determinará la dosis real que será más adecuada para un individuo y variará con la edad, el peso y la respuesta del individuo en particular. Las dosis anteriores son ejemplos del caso promedio. Por supuesto, puede haber casos individuales en los que se merecen intervalos de dosificación más altos o más bajos, y estos están dentro del alcance de esta invención.

10 El intervalo de dosificación requerido depende de la elección del péptido, la ruta de administración, la naturaleza de la formulación, la naturaleza de la condición del sujeto y el juicio del médico tratante. Sin embargo, las dosis adecuadas están en el intervalo de 0,1-100 pg/kg del sujeto.

15 Una composición de la vacuna está convenientemente en forma inyectable. Se pueden emplear adyuvantes convencionales para mejorar la respuesta inmune. Una dosis unitaria adecuada para la vacunación es 0,5-5 microgramos/kg de antígeno, y dicha dosis se administra preferiblemente 1-3 veces y con un intervalo de 1-3 semanas. Con el intervalo de dosis indicado, no se observarán efectos toxicológicos adversos con los compuestos de la invención que impedirían su administración a individuos adecuados.

20 Sin embargo, son de esperar grandes variaciones en la dosificación necesaria en vista de la variedad de compuestos disponibles y las diferentes eficiencias de diversas vías de administración. Por ejemplo, se esperaría que la administración oral requiera dosis más altas que la administración por inyección intravenosa. Las variaciones en estos niveles de dosificación se pueden ajustar usando rutinas empíricas estándar para la optimización, como se entiende bien en la técnica.

25 Bases de datos de secuencias, secuencias en un medio tangible y algoritmos

30 Las secuencias de polinucleótidos y polipéptidos forman un valioso recurso de información con el que determinar sus estructuras bidimensionales y tridimensionales, así como para identificar secuencias adicionales de homología similar. Estos enfoques se facilitan más fácilmente almacenando la secuencia en un medio legible por ordenador y luego utilizando los datos almacenados en un programa conocido de estructura macromolecular o para buscar en una base de datos de secuencia utilizando herramientas de búsqueda bien conocidas, tales como el paquete del programa GCG.

35 También se divulgan métodos para el análisis de secuencias de caracteres o cadenas, particularmente secuencias genéticas o secuencias de proteínas codificadas. Los métodos preferidos de análisis de secuencia incluyen, por ejemplo, métodos de análisis de homología de secuencia, tales como análisis de identidad y similitud, análisis de ADN, ARN y estructura de proteínas, ensamblaje de secuencias, análisis por cladística, análisis de motivo de secuencia, determinación de marco de lectura abierto, llamada de base de ácido nucleico, análisis de uso de codones, recorte de bases de ácido nucleico y análisis de pico de cromatograma de secuenciación.

40 Se proporciona un método basado en ordenador para realizar la identificación de homología. Este método comprende las etapas de: proporcionar una primera secuencia de polinucleótidos que comprende la secuencia de un polinucleótido de la invención en un medio legible por ordenador; y comparar dicha primera secuencia de polinucleótidos con al menos una segunda secuencia de polinucleótidos o polipéptidos para identificar la homología.

45 También se proporciona un método basado en ordenador para realizar la identificación de homología, dicho método comprende las etapas de: proporcionar una primera secuencia de polipéptidos que comprende la secuencia de un polipéptido de la divulgación en un medio legible por ordenador; y comparar dicha primera secuencia de polipéptidos con al menos un segundo polinucleótido o secuencia de polipéptidos para identificar la homología.

50 Definiciones

55 "Identidad", como se conoce en la técnica, es una relación entre dos o más secuencias de polipéptidos o dos o más secuencias de polinucleótidos, según sea el caso, como se determina comparando las secuencias. En la técnica, "identidad" también significa el grado de relación de secuencia entre secuencias de polipéptidos o polinucleótidos, según sea el caso, de acuerdo con lo determinado por la coincidencia entre cadenas de tales secuencias. La "identidad" se puede calcular fácilmente mediante métodos conocidos, que incluyen, entre otros, los descritos en (Computational Molecular Biology, Lesk, AM, ed., Oxford University Press, Nueva York, 1988; Biocomputing: Informatics and Genome Projects, Smith, DW, ed., Academic Press, Nueva York, 1993; Computer Analysis of Sequence Data, Parte I, Griffin, AM, y Griffin, HG, eds., Humana Press, New Jersey, 1994; Sequence Analysis in Molecular Biology, von Heine, G., Academic Press, 1987; y Sequence Analysis Primer, Gribskov, M. y Devereux, J., eds., M Stockton Press, Nueva York, 1991; y Carillo, H., y Lipman, D., SIAM J. Applied Math., 48: 1073 (1988). Los métodos para determinar la identidad están diseñados para proporcionar la mayor coincidencia entre las secuencias probadas. Además, los métodos para determinar la identidad están codificados en programas de ordenador disponibles al público. Los métodos de programas de ordenador para determinar la identidad entre dos secuencias incluyen, entre otros, el programa GAP en el paquete del programa GCG (Devereux, J., et al., Nucleic Acids Research

12 (1): 387 (1984)), BLASTP, BLASTN (Altschul, S.F. et al., J. Molec. Biol. 215: 403-410 (1990), y FASTA (Pearson y Lipman Proc. Natl. Acad. Sci. USA 85: 2444-2448 (1988). La familia de programas BLAST está disponible públicamente a través del NCBI y otras fuentes (BLAST Manual, Altschul, S., et al., NCBI NLM NIH Bethesda, MD 20894; Altschul, S., et al., J. Mol. Biol. 215: 403-410 (1990). El conocido algoritmo de Smith Waterman también puede usarse para determinar la identidad.

Los parámetros para la comparación de secuencias de polipéptidos incluyen lo siguiente:

Algoritmo: Needleman y Wunsch, J. Mol Biol. 48: 443-453 (1970)

Matriz de comparación: BLOSSUM62 de Henikoff y Henikoff, Proc. Natl. Acad. Sci. USA. 89: 10915-10919 (1992)

Penalización por hueco: 8

Penalización por longitud de hueco: 2

Un programa útil con estos parámetros está disponible públicamente como el programa "gap" de Genetics Computer Group, Madison WI. Los parámetros antes mencionados son los parámetros predeterminados para las comparaciones de péptidos (junto con ninguna penalización para los huecos en los extremos).

Los parámetros para la comparación de polinucleótidos incluyen lo siguiente:

Algoritmo: Needleman y Wunsch, J. Mol Biol. 48: 443-453 (1970)

Matriz de comparación: coincidencias = +10, faltas de coincidencia = 0

Penalización por hueco: 50

Penalización por longitud de hueco: 3

Disponible como: El programa "gap" de Genetics Computer Group, Madison WI. Estos son los parámetros predeterminados para las comparaciones de ácidos nucleicos.

Un significado preferido para "identidad" para polinucleótidos y polipéptidos, según sea el caso, se proporciona en (1) y (2) a continuación.

(1) Las realizaciones de polinucleótidos incluyen además un polinucleótido aislado que comprende una secuencia de polinucleótidos que tiene al menos un 50, 60, 70, 80, 85, 90, 95, 97 o 100% de identidad con la secuencia de referencia de la SEQ ID NO: 11, en la que dicha secuencia de polinucleótidos puede ser idéntica a la secuencia de referencia de la SEQ ID NO: 11 o puede incluir hasta un cierto número entero de alteraciones de nucleótidos en comparación con la secuencia de referencia, en la que dichas alteraciones se seleccionan del grupo que consiste en al menos una eliminación, sustitución, incluida transición y transversión, o inserción de nucleótidos, y en la que dichas alteraciones pueden ocurrir en las posiciones terminales 5' o 3' de la secuencia de nucleótidos de referencia o en cualquier lugar entre esas posiciones terminales, intercaladas individualmente entre los nucleótidos en la secuencia de referencia o en uno o más grupos contiguos dentro de la secuencia de referencia, y en la que dicho número de alteraciones de nucleótidos se determina multiplicando el número total de nucleótidos en la SEQ ID NO: 11 por el número entero que define el porcentaje de identidad dividido por 100 y luego restando ese producto de dicho número total de nucleótidos en la SEQ ID NO: 11, o:

$$n_n \leq x_n - (x_n \bullet y),$$

en la que n_n es el número de alteraciones de nucleótidos, x_n es el número total de nucleótidos en la SEQ ID NO: 11, y es 0,50 para 50%, 0,60 para 60%, 0,70 para 70%, 0,80 para 80%, 0,85 para 85%, 0,90 para 90%, 0,95 para 95%, 0,97 para 97% o 1,00 para 100%, y \bullet es el símbolo para el operador de multiplicación, y en el que cualquier producto no entero de x_n e y se redondea hacia abajo al entero más cercano antes de restarlo de x_n . Las alteraciones de las secuencias de polinucleótidos que codifican los polipéptidos de la SEQ ID NO: 1-2 y 4-10 pueden crear mutaciones sin sentido, carentes de sentido o de cambio de marco en esta secuencia de codificación y, por lo tanto, alterar el polipéptido codificado por el polinucleótido después de tales alteraciones.

A modo de ejemplo, una secuencia de polinucleótidos de la presente divulgación puede ser idéntica a las secuencias de referencia de la SEQ ID NO: 11, es decir, puede ser 100% idéntica o puede incluir hasta un cierto número entero de alteraciones de ácido nucleico en comparación con una secuencia de referencia de modo que el porcentaje de identidad sea inferior al 100% de identidad. Dichas alteraciones se seleccionan del grupo que consiste en al menos una eliminación, sustitución de ácido nucleico, incluida la transición y transversión, o inserción de ácidos nucleicos, y en las que dichas alteraciones pueden ocurrir en las posiciones terminales 5' o 3' de la secuencia de polinucleótidos de referencia o en cualquier lugar entre esas posiciones terminales, intercaladas individualmente entre los ácidos nucleicos en la secuencia de referencia o en uno o más grupos contiguos dentro de la secuencia de referencia. El número de alteraciones de ácido nucleico para un porcentaje de identidad dado se determina multiplicando el número total de ácidos nucleicos en la SEQ ID NO: 11 por el número entero que define el porcentaje de identidad dividido por 100 y luego restando ese producto de dicho número total de ácidos nucleicos en la SEQ ID NO: 11, o:

$$n_n \leq x_n - (x_n \bullet y),$$

en la que n_n es el número de alteraciones de ácido nucleico, x_n es el número total de ácidos nucleicos en la SEQ ID NO: 11, y es, por ejemplo, 0,70 para 70%, 0,80 para 80%, 0,85 para 85% etc., \bullet es el símbolo para el operador de

multiplicación, y en la que cualquier producto no entero de x_n y se redondea hacia abajo al entero más cercano antes de restarlo de x_n .

(2) Las realizaciones de polipéptidos incluyen además un polipéptido aislado que comprende un polipéptido que tiene al menos una identidad del 50, 60, 70, 80, 85, 90, 95, 97 o 100% con la secuencia de referencia del polipéptido de la SEQ ID NO: 1-2 y 4-10, en la que dicha secuencia de polipéptidos puede ser idéntica a la secuencia de referencia de la SEQ ID NO: 1-2 y 4-10 o puede incluir hasta un cierto número entero de alteraciones de aminoácidos en comparación con la secuencia de referencia, en la que dichas alteraciones se seleccionan del grupo que consiste en al menos una eliminación, sustitución de aminoácidos, que incluye sustitución o inserción conservadora y no conservadora, y en la que dichas alteraciones pueden ocurrir en las posiciones del terminal amino o carboxilo de la secuencia del polipéptido de referencia o en cualquier lugar entre esas posiciones terminales, intercaladas individualmente entre los aminoácidos en la secuencia de referencia o en uno o más grupos contiguos dentro de la secuencia de referencia, y en la que se determina dicho número de alteraciones de aminoácidos multiplicando el número total de aminoácidos en la SEQ ID NO: 1-2 y 4-10 por el número entero que define el porcentaje de identidad dividido por 100 y luego restando ese producto de dicho número total de aminoácidos en la SEQ ID NO: 1-2 y 4-10, respectivamente, o:

$$n_a \leq x_a - (x_a \cdot y),$$

en la que n_a es el número de alteraciones de aminoácidos, x_a es el número total de aminoácidos en la SEQ ID NO: 1-2 y 4-10, y es 0,50 para 50%, 0,60 para 60%, 0,70 para 70%, 0,80 para 80%, 0,85 para 85%, 0,90 para 90%, 0,95 para 95%, 0,97 para 97% o 1,00 para 100%, y \bullet es el símbolo para el operador de multiplicación, y en la que cualquier producto no entero de x_a y y es redondeado al entero más cercano antes de restarlo de x_a .

A modo de ejemplo, una secuencia de polipéptidos de la presente divulgación puede ser idéntica a la secuencia de referencia de la SEQ ID NO: 1-2 y 4-10, es decir, puede ser 100% idéntica o puede incluir hasta cierto número entero de alteraciones de aminoácidos en comparación con la secuencia de referencia, de modo que el porcentaje de identidad sea inferior al 100% de identidad. Dichas alteraciones se seleccionan del grupo que consiste en al menos una eliminación, sustitución de aminoácidos, que incluye sustitución o inserción conservadora y no conservadora, y en la que dichas alteraciones pueden ocurrir en las posiciones del terminal amino o carboxilo de la secuencia del polipéptido de referencia o en cualquier lugar entre esas posiciones terminales, intercaladas individualmente entre los aminoácidos en la secuencia de referencia o en uno o más grupos contiguos dentro de la secuencia de referencia. El número de alteraciones de aminoácidos para un % de identidad dado se determina multiplicando el número total de aminoácidos en la SEQ ID NO: 1-2 y 4-10 por el número entero que define el porcentaje de identidad dividido por 100 y luego restando ese producto de dicho número total de aminoácidos en la SEQ ID NO: 1-2 y 4-10, o:

$$n_a \leq x_a - (x_a \cdot y),$$

en la que n_a es el número de alteraciones de aminoácidos, x_a es el número total de aminoácidos en la SEQ ID NO: 1-2 y 4-10, y es, por ejemplo, 0,70 para 70%, 0,80 para 80%, 0,85 para 85% etc., y \bullet es el símbolo para el operador de multiplicación, y en el que cualquier producto no entero de x_a y y se redondea hacia abajo al entero más cercano antes de restarlo de x_a .

El "Individuo o individuos", cuando se usan en el presente documento con referencia a un organismo, significa un eucariota multicelular, que incluye, pero no se limita a, un metazoo, un mamífero, un óvulo, un bóvido, un simio, un primate y un humano.

"Aislado" significa alterado "por la mano del hombre" a partir de su estado natural, es decir, si ocurre en la naturaleza, se ha cambiado o eliminado de su entorno original, o ambos. Por ejemplo, un polinucleótido o un polipéptido presente de forma natural en un organismo vivo no está "aislado", pero el mismo polinucleótido o polipéptido separado de los materiales coexistentes de su estado natural está "aislado", como se emplea el término en el presente documento. Además, un polinucleótido o polipéptido que se introduce en un organismo por transformación, manipulación genética o por cualquier otro método recombinante se "aisla" incluso si todavía está presente en dicho organismo, organismo que puede estar vivo o no vivo.

"Polinucleótido o polinucleótidos" generalmente se refiere a cualquier polirribonucleótido o polidesoxirribonucleótido, que puede ser ARN o ADN no modificado o ARN o ADN modificado que incluye regiones monocatenarias y bicatenarias.

"Variante" se refiere a un polinucleótido o polipéptido que difiere de un polinucleótido o polipéptido de referencia, pero conserva propiedades esenciales. Una variante típica de un polinucleótido difiere en la secuencia de nucleótidos de otro polinucleótido de referencia. Los cambios en la secuencia de nucleótidos de la variante pueden o no alterar la secuencia de aminoácidos de un polipéptido codificado por el polinucleótido de referencia. Los cambios de nucleótidos pueden dar como resultado sustituciones, adiciones, eliminaciones, fusiones y truncamientos de aminoácidos en el polipéptido codificado por la secuencia de referencia, como se discute a continuación. Una variante típica de un polipéptido difiere en la secuencia de aminoácidos de otro polipéptido de referencia. En general, las diferencias son

limitadas, de modo que las secuencias del polipéptido de referencia y la variante son muy similares en general y, en muchas regiones, idénticas. Una variante y un polipéptido de referencia pueden diferir en la secuencia de aminoácidos en una o más sustituciones, adiciones, eliminaciones en cualquier combinación. Un residuo de aminoácido sustituido o insertado puede o no estar codificado por el código genético. Una variante de un polinucleótido o polipéptido puede ser de origen natural, tal como una variante alélica, o puede ser una variante que no se sabe que se produce de forma natural. Las variantes de polinucleótidos y polipéptidos que no se producen de forma natural pueden prepararse mediante técnicas de mutagénesis o mediante síntesis directa.

"Enfermedad o enfermedades" significa cualquier enfermedad causada o relacionada con una infección por una bacteria, que incluye, por ejemplo, otitis media en bebés y niños, neumonía en ancianos, sinusitis, infecciones nosocomiales y enfermedades invasivas, otitis media crónica con pérdida de audición, acumulación de líquido en el oído medio, daño del nervio auditivo, retraso en el aprendizaje del habla, infección del tracto respiratorio superior e inflamación del oído medio.

Parte experimental

Los ejemplos a continuación se llevan a cabo utilizando técnicas estándar, que son bien conocidas y rutinarias para los expertos en la técnica, excepto donde se describa lo contrario en detalle. Los ejemplos son ilustrativos, pero no limitan la invención.

La presente investigación describe el aislamiento, purificación, caracterización, clonación y expresión de la nueva proteína de membrana externa llamada proteína E (pE) de *H. influenzae* y la nueva pE(A) recombinante truncada, que se descubrió usando un suero de mieloma con IgD(λ) humana.

Materiales y métodos

Reactivos

Las cepas MinnA y NTHi3655 de *H. influenzae* tipo b se obtuvieron amablemente de Robert S. Munson Jr. (Escuela de Medicina de la Universidad de Washington (St. Louis, Mo). La cepa NTHi772 de *H. influenzae* no tipificable era un aislado clínico de un cultivo de hisopo nasofaríngeo en nuestro Departamento (2). También se analizó una serie de diferentes especies de *Haemophilus* y se describen en la Tabla 1. El suero completo de mieloma con IgD humana IgD(λ) se adquirió a través de The Binding Site (Birmingham, Inglaterra). Para producir un antisuero anti-pE específico, los conejos se inmunizaron por vía intramuscular con 200 μ g de pE22-160 recombinante [pE(A)] emulsionado en adyuvante completo de Freund (Difco, Becton Dickinson, Heidelberg, Alemania) o péptido pE41-68 conjugado con hemocianina de lapa ojo de cerradura (KLH) y reforzada los días 18 y 36 con la misma dosis de proteína en adyuvantes incompletos de Freund. Se extrajo sangre de 2 a 3 semanas después. Los anticuerpos policlonales resultantes se aislaron por cromatografía de afinidad usando pE(A) o un péptido pE específico (pE41-68) conjugado con CnBr-Sefarosa (11). El anticuerpo de cabra anti-IgD humano conjugado con peroxidasa de rábano picante (HRP) era de Biosource (Camarillo, CA). pAb de conejo anti-IgD humano era de Dakopatts (Gentofte, Dinamarca).

El anticuerpo de ratón anti-IgD humano conjugado con isotiocianato de fluoresceína (FITC), cadenas ligeras antihumanas de conejo conjugadas con HRP (κ y λ), e inmunoglobulinas policlonales de cerdo anti-conejo conjugadas con FITC se adquirieron a través de Dakopatts.

Extracción y purificación de proteína E

H. influenzae tipo b (MinnA) se cultivó durante la noche en caldo de infusión de cerebro y corazón (BHI) (Laboratorios Difco, Detroit, Michigan) suplementado con NAD y hemina (Sigma, St. Louis, MO), cada uno a 10 μ g/mL. Después de dos lavados, las bacterias se extrajeron en tampón Tris-HCl 0,05 M (pH 8,8) que contenía Empigen® al 0,5% (Calbiochem Novabiochem, Bedford, MA). La suspensión bacteriana se mezcló por agitación magnética durante 2 horas a 37 °C. Después de la centrifugación a 8000 x g durante 20 min a 4 °C, el sobrenadante se filtró con un filtro estéril (0,45 μ m; Sterivex-HV, Millipore). Se aplicó extracto de *H. influenzae* en Empigen® al 0,5% a una columna de Q-sefarosa (Amersham Pharmacia Biotech) equilibrada con Tris-HCl 0,05 M (pH 8,8) que contenía urea 6 M. La columna se eluyó usando un gradiente lineal de NaCl 0 a 1 M en el mismo tampón. Las fracciones detectadas por el suero de mieloma con IgD(λ) se agruparon, se dializaron en tubos de membrana Spectraphor (corte de peso molecular 6-8.000; Spectrum, Laguna hills, CA) contra Tris-HCl 0,05 M, pH 8,8 y se concentraron en membranas de disco YM100 (corte de peso molecular 10.000; Amicon, Beverly, MA).

SDS-PAGE y detección de proteínas en membranas (transferencia Western)

La SDS-PAGE se corrió a un voltaje constante de 150 usando geles Bis-Tris al 10% con tampón de operación (MES), de muestra (LDS) y de transferencia, así como un instrumento de transferencia de Novex (San Diego, CA). Las muestras se calentaron regularmente a 100 °C durante 10 min. Los geles se tiñeron con azul brillante Coomassie R-250 (13; Bio-Rad, Sundbyberg, Suecia). La transferencia electroforética de las bandas de proteínas del gel a una membrana de Immobilon-P (Millipore, Bedford, MA) se realizó a 30 V durante 2 a 3 horas. Después de la transferencia,

la membrana de Immobilon-P se bloqueó en PBS con Tween 20 al 0,05% (PBS-Tween) que contenía leche en polvo al 5%. Después de varios lavados en PBS-Tween, la membrana se incubó con proteína purificada de mieloma con IgD (0,5 µg/mL, mieloma con IgD(λ) humana; The Binding Site) en PBS-Tween incluyendo leche en polvo al 2% durante 1 hora a temperatura ambiente. Se añadió anticuerpo de cabra anti-IgD humano conjugado con HRP diluido 1/1000 después de varios lavados en PBS-Tween. Después de la incubación durante 40 minutos a temperatura ambiente y varios lavados adicionales en PBS-Tween, el desarrollo se realizó con reactivos de detección de transferencia Western ECL (Amersham Pharmacia Biotech, Uppsala, Suecia).

Electroforesis bidimensional en gel de SDS-poliacrilamida (PAGE 2-D) y transferencia Western

Después de la cromatografía de intercambio iónico, los extractos de Empigen® de *H. influenzae* (MinnA) se sometieron a enfoque isoeléctrico (IEF) usando el sistema de IEF IPGphor (Amersham Pharmacia Biotech) (5,12). Para la calibración del gel, se utilizó un estándar (cat. No. 161-0320; Bio-Rad). Los geles de poliácridamida 2-D se transfirieron eléctricamente a filtros Immobilon-PVDF (0,45 mm; Millipore, Bedford, USA.) A 120 mA durante la noche. Después de la saturación, las etapas de incubación, bloqueo y lavado se realizaron como se describió anteriormente.

Análisis de secuencia de aminoácidos

El análisis automático de la secuencia de aminoácidos se realizó con un secuenciador 470A en fase sólida gas-líquido de Applied Biosystems (Foster City, CA).

Construcción de una biblioteca genómica de *H. influenzae*

El ADN cromosómico se preparó a partir de la cepa 772 usando una modificación del método de Berns y Thomas (2,13). En resumen, se construyó una biblioteca genómica de *H. influenzae* 772 a partir de 40 µg de ADN que se digirió parcialmente con Sau3A durante 1 hora. El ADN escindido se fraccionó en un gradiente de sacarosa (14). Se agruparon las fracciones que contenían fragmentos de ADN de tamaños apropiados (2 a 7 kpb), y el ADN se ligó a pUC18 digerido con BamHI seguido de transformación en *Escherichia coli* JM83 por electroporación usando un pulsador de genes (Bio-Rad, Richmond CA). La bacteria se sembró en placas sobre agar LB suplementado con ampicilina y X-Gal (5-bromo-4-cloro-3-indolil-β-D-galactopiranosido).

Inmunoensayo de colonias, aislamiento de ADN y secuenciación

Los transformantes de *E. coli*, cultivados durante la noche en agar LB, se transfirieron a filtros de nitrocelulosa (Sartorius, Göttingen, Alemania) cubriendo las superficies de agar con filtros secos. Las placas se dejaron durante 15 minutos, y las células se lisaron por exposición a vapor de cloroformo saturado durante 20 minutos. Los sitios de unión a proteínas residuales en los filtros se bloquearon incubando los filtros en solución salina equilibrada con Tris que contenía ovoalbúmina (clorhidrato de Tris 50 mM, NaCl 154 mM, ovoalbúmina al 1,5% [pH 7,4]) durante 30 minutos. Después de ser bloqueados, los filtros se incubaron con IgD(λ) humana durante 30 minutos. Se agregaron anticuerpos policlonales anti-IgD humanos conjugados con HRP después de los lavados y los filtros se incubaron durante 30 minutos. Todas las incubaciones se realizaron a temperatura ambiente. Finalmente, los filtros se desarrollaron utilizando 4-cloro-1-naftol y H₂O₂. Se recogieron clones positivos y se purificó el ADN plasmídico pUC18 que contenía ADN genómico de NTHi772. El inserto de ADN de NTHi772 resultante se secuenció usando los cebadores flanqueantes M13+ y M13- y el kit de secuenciación de reacción BigDye Terminator Cycle Sequencing v. 2.0 Ready (Perkin-Elmer, Foster City, Ca). La secuencia obtenida del inserto (3,55 kpb) correspondió a un tramo que contenía ADN que codifica las proteínas HI0175, HI0176, HI0177 y finalmente HI0178 (15).

Clonación de ADN y expresión de proteínas.

Todos los constructos se fabricaron usando fragmentos amplificados por PCR. Se usó pUC18 que contenía ADN genómico de NTHi 772 (HI0175 a HI0178) como plantilla. La ADN polimerasa Taq era de Roche (Mannheim, Alemania) y las condiciones de PCR fueron las recomendadas por el fabricante. Se clonaron los marcos de lectura abiertos de las cuatro proteínas predictivas HI0175 a HI0178, pero solo se ha incluido en el presente documento el procedimiento que describe la clonación de HID (HI0178). pE²²⁻¹⁶⁰ [designado pE(A)] estaba desprovisto del péptido señal endógeno que incluía el residuo de aminoácido glutamina²¹ y se amplificó por PCR usando los cebadores 5'-ctcaggatccaaaggctgaacaaaatgatgtg-3' (SEQ ID NO: 13) y 5'-gggtcgagattaagctttttatcaactg-3' (SEQ ID NO: 14) que introduce los sitios de enzimas de restricción *Bam*HI y *Hind*III. Para fusionar 6 residuos de histidina (SEQ ID NO: 12) codificados por el vector de expresión, se mutó el codón de parada de pE²²⁻¹⁶⁰. El marco de lectura abierto resultante de 417 pb del gen *pe* se ligó a pET26(+) (Novagen, Darmstadt, Alemania). Para evitar la presunta toxicidad, los plásmidos resultantes se transformaron primero en el huésped *E. coli* DH5α. Posteriormente, los plásmidos que codifican pE y pE(A) se transformaron en el huésped de expresión BL21 (DE3). Además de pE y pE(A) de longitud completa, se fabricaron una serie de variantes de pE truncadas. Un esquema se muestra en la Figura 8. Los cebadores que contienen *Bam*HI y *Hind*III se usaron para todos los constructos. Los procedimientos para las variantes truncadas fueron como se describió anteriormente. Todos los constructos se secuenciaron utilizando el kit de secuenciación de reacción BigDye Terminator Cycle Sequencing v. 2.0 Ready (Perkin-Elmer, Foster City, Ca).

Para producir proteínas recombinantes, las bacterias se cultivaron hasta la fase media logarítmica (OD_{600} de 0,6 a 0,8) seguido de inducción con isopropil-1-tio- β -D-galactósido (IPTG) 1 mM, dando como resultado una sobreexpresión de pE(A). Cuando la OD_{600} alcanzó 1,5 a 1,7, se cosecharon las bacterias y se aislaron los cuerpos de inclusión de acuerdo con un protocolo estándar. Las proteínas recombinantes podrían purificarse adicionalmente por cromatografía de afinidad usando columnas de níquel. Las proteínas recombinantes purificadas se analizaron posteriormente por SDS-PAGE.

Purificación de ADN de *H. influenzae*, condiciones de PCR y secuenciación

Se aisló ADN genómico de aislados clínicos de *H. influenzae* usando un kit de tejido DNeasy (Qiagen, Hilden, Alemania). La ADN polimerasa Taq era de Roche (Mannheim, Alemania) y las condiciones de PCR fueron las recomendadas por el fabricante. Para aislar el gen *pe*, se usaron el par de cebadores 5'-gcattattaggctgagttattg-3' (SEQ ID NO: 15) y 5'-gaaggattattctgatgcag-3' (SEQ ID NO: 16), que hibridan con los genes de flanco HI0177, HI0179 respectivamente. Los productos de PCR resultantes (948 pb) se secuenciaron mediante desplazamiento sobre el genoma usando los cebadores mencionados anteriormente además de los cebadores 5'-cttgggttactaccgcttg-3' (SEQ ID NO: 17) y 5'-gtgttaaactaacgtatg-3' (SEQ ID NO: 18). La electroforesis capilar se realizó en un Beckman CEQ 2000 utilizando un kit de secuenciación de ciclo de terminación de coloración (el kit de DTCS CEQ, Beckman Coulter, Estocolmo, Suecia). La edición y la alineación de las secuencias de ADN resultantes se realizaron usando PHRED (CodonCode, Deadham, EE. UU.) Y SEQUENCHER (MedProbe, Oslo, Noruega).

Fabricación de una *H. influenzae* deficiente en *pe* (NTHi 3655 Δ *pe*)

Se usó ADN genómico aislado de NTHi 772 como plantilla. Los extremos 5' y 3' de *pe*, incluidas partes de los genes HI0177 y HI0179, se amplificaron como dos casetes (815 pb y 836 pb, respectivamente) utilizando ADN polimerasa DyNAzyme^{MR} II (Finnzymes, Espoo, Finlandia) que introduce los sitios de las enzimas de restricción *Xho*I y *Eco*RI respectivamente *Eco*RI y *Sp*EI además de secuencias de captación específicas en los dos casetes (18). Los fragmentos de PCR resultantes se digirieron y clonaron en pBluescript SK (+/-). Se obtuvo un casete del gen de resistencia a la kanamicina (1282 pb) a partir de pUC4K usando el sitio de enzima de restricción para *Eco*RI. Después de la digestión, el producto de PCR se ligó al fragmento del gen *pe* truncado que contenía partes de los genes HI0177 y HI0179. Las cepas Eagan y RM804 de *H. influenzae* se transformaron de acuerdo con el método M-IV de Heriott et al. (19). Los mutantes resultantes se verificaron mediante PCR y la expresión de *pe* se analizó mediante transferencia Western y citometría de flujo.

Softwares de biología molecular

Las secuencias obtenidas se compararon con el genoma KW20 disponible de *H. influenzae* (<http://www.tigr.org>) (15). El péptido señal se dedujo utilizando el Centro de Servidores de Predicción World Wide Web SignalP V1.1 para el Análisis de Secuencias Biológicas (<http://www.cbs.dtu.dk/services/SignalP/>) (16). El perfil de hidrofobicidad de *pe* se analizó mediante el método de Kyte y Doolittle (17).

Animales, procedimientos quirúrgicos y modelo de otitis media en ratas.

Se usaron ratas Sprague-Dawley macho sanas, con un peso de 200-250 g. Todos los animales estaban libres de infecciones del oído medio de acuerdo con lo determinado por otomicroscopía antes de la operación. En las intervenciones, las ratas se anestesiaron con metohexital (Brietal®, Elli Lilly and Company, Indianápolis, IN) por vía intravenosa o hidrato de cloral (apoteksbolaget, Lund, Suecia) por vía intraperitoneal. Se cultivaron bacterias para experimentos con animales como se describió anteriormente. Después de la cosecha por centrifugación, las bacterias se resuspendieron en medio de cultivo fresco a una concentración de 2×10^{10} unidades formadoras de colonias (ufc) tanto para NTHi 3655 como para el mutante *pe* correspondiente. Las preparaciones se mantuvieron en hielo hasta su uso. Para inducir otitis media aguda (OMA), se alcanzó el oído medio mediante una incisión ventral en la línea media del cuello y se instilaron aproximadamente 0,05 mL de la suspensión bacteriana en la cavidad del oído medio. La otomicroscopía se realizó los días 3 y 5 después de la operación. La otitis media con efusión purulenta, es decir, una efusión opaca y una dilatación a menudo pronunciada de los vasos en el día 3, se denominó OMA.

Resultados

Extracción y separación de una proteína (*pe*) de *H. influenzae* que es detectada por un suero de mieloma con IgD(λ)

Se había pensado anteriormente que *H. influenzae* no tipificable (NTHi) no se une a IgD (19), mientras que *H. influenzae* encapsulada se une fuertemente a IgD. Se descubrió que un suero de mieloma con IgD(λ) se unía específicamente también a NTHi. Un perfil típico de citometría de flujo de la NTHi772 se muestra en la Figura 1. Se encontró un cambio fuerte y un aumento de la fluorescencia en presencia de la proteína de mieloma con IgD(λ) en comparación con el control sin IgD.

Para analizar en detalle la proteína de la membrana externa de *H. influenzae* que fue detectada por el mieloma con IgD(λ), se solubilizó una fracción de la membrana externa en el detergente Empigen®. La Figura 1B demuestra que

se obtuvo una actividad de unión a IgD muy fuerte en transferencias Western para una proteína con un peso molecular aparente de aproximadamente 16 kDa. Sin embargo, no se pudo detectar una banda de proteína distinta correspondiente a la actividad de unión a IgD en la SDS-PAGE teñida de azul brillante Coomassie. Después de la separación en una columna de Q-Sefarosa, se aplicó el mismo extracto de membrana externa a electroforesis en gel bidimensional y tinción con plata (Figura 1C). Paralelamente, se realizó una transferencia Western correspondiente sondeada con IgD(λ) humana. Por lo tanto, el área donde se localizó pE podría ser encerrada. Sin embargo, no se observó proteína visible (Figura 1C).

Clonación de la proteína E y fabricación de un mutante de *H. influenzae* no tipificable desprovisto de pE

Dado que no se pudo detectar ninguna proteína en el análisis 2-D después de la separación, se construyó una biblioteca de ADN de *H. influenzae* usando la cepa 772 de *H. influenzae* no tipificable (NTHi). Fragmentos que contienen ADN genómico en el intervalo de 2 a 7 kpb se ligaron a pUC18 seguido de transformación en *E. coli* JM83. Los transformantes se analizaron para la unión a IgD usando un inmunoensayo de colonias que consiste en IgD(λ) humanos y anticuerpos policlonales anti-IgD humanos conjugados con HRP. Se encontraron tres colonias positivas de las 20.000 colonias analizadas y se las sometió a una segunda ronda de detección con IgD(λ). Se secuenció uno de los clones positivos y se encontró un inserto de 3,55 kb que contiene ADN que codifica las cuatro proteínas HI0175 a HI0178 de acuerdo con el mapa físico de *H. influenzae* KW20 (15). Para verificar aún más la interacción específica con IgD(λ), el transformante seleccionado también se analizó por citometría de flujo. Como se puede observar en la Figura 2B, *E. coli* JM83 que alberga el ADN genómico de *H. influenzae* 772 correspondiente a la secuencia que codifica HI0175 a HI0178 fue detectado para IgD(λ) en comparación con *E. coli* de control negativo transformada con un vector vacío únicamente (Figura 2A).

Además del análisis del clon JM83 de *E. coli*, las cuatro proteínas de *H. influenzae* (HI0175 a HI0178) se clonaron en el vector de expresión pET26(+) y se produjeron en *E. coli* BL21DE3. Las proteínas recombinantes resultantes se analizaron por IgD(λ) en transferencias Western y se encontró que HI0178 era la única proteína que se detectó por IgD(λ) (datos no mostrados).

Se mutó una *H. influenzae* no tipificable (NTHi 3655) mediante la introducción de un casete del gen de resistencia a la kanamicina en el gen que codifica para pE. Los mutantes resultantes se confirmaron mediante PCR y la ausencia de la expresión de pE se probó mediante el análisis de proteínas de la membrana externa en transferencias Western usando un antisuero específico anti-pE (datos no mostrados). El mutante NTHi 3655 Δ pE también se probó por citometría de flujo y se encontró una fluorescencia claramente disminuida con el mutante en comparación con el tipo silvestre NTHi 3655 correspondiente cuando se analizó con un antisuero monovalente de conejo anti-pE y un pAb secundario de cabra anti-conejo conjugado con FITC (Figura 2C y D).

La proteína E se detectó en todas las *H. influenzae*, mientras que otras subespecies fueron negativas

Para analizar la expresión de pE de aislados clínicos y cepas tipo de NTHi, se desarrolló un ensayo de citometría de flujo de unión directa que consiste en el suero con IgD(λ) y un anticuerpo secundario conjugado con FITC dirigido contra IgD humana. En los experimentos iniciales, las bacterias recolectadas en diferentes puntos de tiempo se analizaron para la expresión de pE. No se observó diferencia con respecto a la expresión de pE entre las bacterias de crecimiento logarítmico o de fase estacionaria, lo que sugiere que la pE de la superficie de NTHi no dependía de la fase de crecimiento. Las bacterias de fase estacionaria se utilizaron así en todos los análisis adicionales. Se analizó la intensidad de fluorescencia media (mfi) por célula bacteriana y se incluyeron un total de 22 cepas de NTHi en este estudio. La intensidad de fluorescencia varió entre diferentes cepas de NTHi, aunque se detectó pE en la mayoría de NTHi en este ensayo particular usando el suero de mieloma con IgD(λ) como anticuerpo de detección (Figura 3).

En otros experimentos, se usaron anticuerpos específicos de conejo anti-pE para la detección de pE expuesta en la superficie. El antisuero específico anti-pE específicamente reconoció pE también en cepas encapsuladas de *H. influenzae*, *H. aegypticus* y *H. haemolyticus* (datos no mostrados).

Para analizar adicionalmente los niveles de expresión de pE, los extractos de membrana externa tratados con Empigen® de varias especies de *Haemophilus* se probaron en transferencias Western usando IgD(λ) como anticuerpo de detección (Tabla 1). En estos experimentos, no se encontraron diferencias entre las cepas de *Haemophilus* de alta y baja expresión, es decir, en las transferencias Western, todas las cepas mostraron pE de la misma intensidad y en la misma posición correspondiente a 16 kDa. *H. influenzae* encapsulada (tipo a hasta f) también expresaron pE como lo revelaron las transferencias Western (Tabla 1). Por ejemplo, la expresión de pE en cuatro cápsulas de *H. influenzae* tipo b (Hib) se compara con NTHi en la Figura 4. *H. aegypticus* y *H. haemolyticus* expresaron pE (Tabla 1), mientras que para otras especies de *Haemophilus* relacionadas, que fueron negativas en citometría de flujo (Figura 3), no se detectó pE en los análisis de transferencia Western. El antisuero anti-pE específico reconoció específicamente pE también en cepas encapsuladas (datos no mostrados).

ES 2 809 167 T3

Tabla 1

	Encapsulada o no tipificable (NT)	Transferencia Western (positiva/negativa; 0)
556	<i>H. influenzae</i> CCUG EF 6881 I Caps.tipo a	pos
557	<i>H. influenzae</i> CCUG EF 7315 I Caps.tipo a	pos
94	H.i. tipo a	pos
479	<i>H. influenzae</i> MinnA tipo b	pos
547	<i>H. influenzae</i> Eagan tipo b	pos
485	<i>H. influenzae</i> 85 05 30 b	pos
539	<i>H. influenzae</i> D-22 Caps.tipo b	pos
541	<i>H. influenzae</i> HK 695 Caps.tipo b	pos
542	<i>H. influenzae</i> HK 691 Caps.tipo b	pos
577	<i>H. influenzae</i> HK 713 Caps.tipo b	pos
578	<i>H. influenzae</i> HK 714 Caps.tipo b	pos
582	<i>H. influenzae</i> HK 83458 caps tipo b	pos
581	<i>H. influenzae</i> HK 163 caps tipo b	pos
580	<i>H. influenzae</i> HK 729 caps tipo b	pos
569	<i>H. influenzae</i> 17 B Dallas caps tipo b	pos
568	<i>H. influenzae</i> DL 42/2F4 caps tipo b	pos
579	<i>H. influenzae</i> HK 720 caps tipo b	pos
477	<i>H. influenzae</i> 6-460 caps tipo b	pos
570	<i>H. influenzae</i> DL 42 caps tipo b	pos
	<i>H. influenzae</i> RM 804 no caps tipo b	pos
583	<i>H. influenzae</i> HK 705 no caps tipo b	pos
551	<i>H. influenzae</i> CCUG EF 4851 II Caps.tipo c	pos
552	<i>H. influenzae</i> CCUG EF 4852 II Caps.tipo c	pos
95	H.i. tipo c	pos
555	<i>H. influenzae</i> CCUG EF 6878 IV Caps.tipo d	pos
560	<i>H. influenzae</i> CCUG EF NCTC 8470 Caps.tipo d	pos
96	H.i. tipo d	pos
554	<i>H. influenzae</i> CCUG EF 6877 IV Caps.tipo e	pos
88	H.i. tipo e A11/01	pos
89	H.i. tipo e A76/01	pos
90	H.i. tipo e A77/99	pos
559	<i>H. influenzae</i> CCUG EF 15519 II Caps.tipo f	pos
78	H.i. CCUG 15435 Caps.tipo f	pos
91	H.i. tipo f A1/01	pos
92	H.i. tipo f A58/01	pos
93	H.i. tipo f A91/01	pos
67	H.i. NT 6-9547 b.tipo l	pos
478	<i>H. influenzae</i> 6-601 NT	pos
484	<i>H. influenzae</i> 6-6200 NT NT	pos

ES 2 809 167 T3

(continuación)

	Encapsulada o no tipificable (NT)	Transferencia Western (positiva/negativa; 0)
507	<i>H. influenzae</i> 6-102 NT NT	pos
65	H.i.NT 7-68/99 Claren. lavadura	pos
68	H.i. NT 7-758	pos
546	<i>H. influenzae</i> 6-9547 NT	pos
480	<i>H. influenzae</i> 772 NT	pos
476	<i>H. influenzae</i> 6-115 NT	pos
481	<i>H. influenzae</i> 6-504 NT	pos
482	<i>H. influenzae</i> 6-7702 NT	pos
483	<i>H. influenzae</i> 82 10 23 NT	pos
486	<i>H. influenzae</i> 6-121 NT	pos
506	<i>H. influenzae</i> 6-6267 NT	pos
540	<i>H. influenzae</i> D-26 NT	pos
543	<i>H. influenzae</i> Buffalo 1479 NT	pos
544	<i>H. influenzae</i> Buffalo C 7961 NT	pos
64	H.i. 56-2934 000428 NT	pos
69	H.i. 7-120/99 bronquios NT	pos
70	H.i. 7-161/99 bronquios NT	pos
66	H.i. S6-2952 000428 NT	pos
105	H.i. RM 3655 NT	pos
531	<i>H. aegypticus</i> EF 628 NT, no b	pos
73	<i>H. aegypticus</i> CCUG 25716	pos
74	<i>H. aegypticus</i> CCUG 26840	pos
76	<i>H. aegypticus</i> CCUG 39154	pos
79	<i>H. aegypticus</i> HK 1247	pos
80	<i>H. aegypticus</i> HK 1229	pos
81	<i>H. aegypticus</i> HK 1242	pos
82	H.i. biovar <i>aegypticus</i> HK 865	pos
83	H.i. biovar <i>aegypticus</i> HK 871	pos
84	H.i. biovar <i>aegypticus</i> HK 1222	pos
85	H.i. biovar <i>aegypticus</i> HK 1239	pos
86	Enfermedad similar a BPF HK 1212	pos
87	Enfermedad similar a BPF HK 1213	pos
72	<i>H. haemolyticus</i> CCUG 15642	pos
101	<i>H. haemolyticus</i> 34669/83	pos
102	<i>H. haemolyticus</i> 47802/88	pos
103	<i>H. haemolyticus</i> 937016	pos
104	<i>H. haemolyticus</i> 74108/81	pos
520	<i>H. parainfluenzae</i> Biotipo I HK 409	0
521	<i>H. parainfluenzae</i> Biotipo II HK 23	0

(continuación)

	Encapsulada o no tipificable (NT)	Transferencia Western (positiva/negativa; 0)
58	59004 H. parainfl.biov.III	0
59	75834 H. parainfl.biov.III	0
60	78909 H. parainfl.biov.III	0
97	<i>H. parainfluenzae</i> 947172	0
98	<i>H. parainfluenzae</i> 59257/91	0
99	<i>H. parainfluenzae</i> 59004/91	0
100	<i>H. parainfluenzae</i> 977101	0
545	<i>H. parainfluenzae</i> Buffalo 3198 NT	0
75	<i>H. som nus</i> CCUG 37617	0
512	<i>H. parasuis</i> 9918	0
516	<i>H. parasuis</i> 99 19	0
527	<i>H. parasuis</i> EF 3712	0
524	<i>H. paraphrophilus</i> HK 319	0
525	<i>H. paraphrophilus</i> HK 415	0
517	<i>H. paraphrophilus</i> 12894	0
71	<i>H. segnis</i> CCUG 14834	0
526	<i>H. aphrophilus</i> HK 327	0
529	<i>H. aphrophilus</i> 11832 A	0
532	<i>H. pleurapneumoniae</i> EF 9917	0
61	39612 <i>Eikenella</i> corrodens	0
62	49064 <i>Eikenella</i> corrodens	0
63	959074 <i>Eikenella</i> corrodens	0
537	<i>Actinobacillus actinomyc.</i> HK 666 <i>P.mult.</i> 78908/90	0

pE22-160 producida de forma recombinante [pE(A)] se detecta por IgD(λ)

- 5 En los experimentos iniciales, se intentó fabricar pE en forma recombinante, pero solo se obtuvieron concentraciones mínimas. Para maximizar el rendimiento de la proteína, se construyó un fragmento de pE truncado que consiste en residuos de aminoácidos lisina22 a lisina160. El péptido señal del terminal N que incluye el aminoácido glutamina21 se eliminó por lo tanto y se reemplazó con el péptido líder además de nueve residuos procedentes del vector pET26(+) (Figura 5A). La pE22-160 truncada se designó pE(A) (Figura 5B). Para confirmar que el producto de proteína recombinante correspondía a pE de 'tipo silvestre', la pE(A) expresada de forma recombinante junto con pE aislada de NTHi 772 se analizó mediante SDS-PAGE y transferencia Western usando la IgD(λ) como sonda. Como se muestra en la Figura 5D, la pE(A) recombinante correspondía significativamente a pE de tipo silvestre en SDS-PAGE. Además, el antisuero con IgD(λ) podría detectar pE(A) producida en *E. coli* hasta 0,01 microgramos (Figura 5E).
- 10
- 15 pE(A) se usó para la inmunización de conejos y después de completar el programa de inmunización como se describe en detalle en Materiales y métodos, el antisuero anti-pE se purificó en una columna que consiste en un péptido inmunodominante calculado (aminoácidos pE41- 68). Los anticuerpos resultantes detectaron claramente pE tanto en la superficie bacteriana (Figura 2C) como en las transferencias Western (no mostradas).
- 20 La secuencia de ADN del gen e de la proteína y el marco de lectura abierto
- El ADN y la secuencia de aminoácidos de pE1-160 de la cepa NTHi 772 se divulga en la Figura 6. El marco de lectura abierto (ORF) tiene una longitud de 160 aminoácidos y el péptido señal predicho tiene una longitud de 20 aminoácidos. El análisis por ordenador sugirió que la señal peptídica reconoce los residuos de aminoácidos alanina18 a lisina22 y se divide entre los residuos isoleucina20 y glutamina21 (38). Paralelamente, el perfil de hidropatía de HID (39) muestra que pE tiene un péptido señal hidrófobo, mientras que el resto de la molécula es principalmente hidrófila (Figura 7).
- 25

Para determinar en detalle el sitio de escisión de la peptidasa señal, la pE recombinante de longitud completa se sometió a degradación de Edman. Sin embargo, el extremo terminal amino de la cadena del polipéptido pE probablemente estaba bloqueada. Una posible explicación para esta falla podría ser que el primer aminoácido era un residuo de piroglutamilo como se describió previamente para la familia de proteínas de *M. catarrhalis* UspA (11). Sin embargo, los intentos de eliminar este supuesto residuo con piroglutamato aminopeptidasa fracasaron. En contraste con pE1-160 de longitud completa, la secuencia del terminal N para pE(A) (pE22-160) que carece del péptido señal de *H. influenzae* endógeno (Figura 5), se caracterizó con éxito por la degradación de Edman y se encontró que contenía la secuencia pronosticada del vector, es decir, la peptidasa señal en *E. coli* escindida en la posición correcta (Figura 5A).

La proteína E fue secuenciada en una serie de diferentes especies de *Haemophilus* incluyendo NTHi y *H. influenzae* encapsulada (Tabla 1). Curiosamente, pE estaba extraordinariamente conservada. Solo unos pocos aminoácidos fueron mutados puntualmente y estos fueron mutados en la mayoría de las cepas siguiendo un patrón específico (Figura 6 y Tabla 2).

Tabla 2. Mutaciones puntuales en el gen pe en diferentes especies de *Haemophilus* en comparación con *H. influenzae* Rd como cepa de referencia^{a)}

Mutación puntual	Número de cepas con mutaciones (%)	Número analizado
Ile20 > Thr20	18 (58 %)	31
Ala23 > Val23	5 (16 %)	31
Glu24 > Lys24	1 ^{b)} (3,2 %)	31
Val28 > Met28	2 (6,5 %)	31
Ala31 > Thr31	19 (61 %)	31
Pro32 > Val32	3 (9,7 %)	31
Pro32 > Ala32	5 (16 %)	31
Ile41 > Val41	8 (26 %)	31
Val47 > Ala47	3 (10 %)	30
Arg76 > Lys76	2 (7,7 %)	26
Ile107 > Val107	16 (62 %)	26
Lys153 > Glu153	1 ^{b)} (3,2 %)	13
Ala154 > Val154	2 (15 %)	13
Ile20 > Thr20	18 (58 %)	31
Terminal C prolongado (3 aa extra)		
SVDKK parada (=Rd) ^{c)} (SEQ ID NO: 19)	11 (85 %)	13
SVDKKSAP parada (SEQ ID NO: 20)	2 (15 %)	13

a) El gen pe fue secuenciado en *H. influenzae* encapsulada tipo a (n=2), b (n=2), c (n=2), d (n=1), e (n=2) y f (n=3), NTHi (n=8), *H. influenzae* biovar *aegypticus* (n=6) y *H. aegypticus* (n=5), usando cebadores flanqueantes.

b) NTHi 772 (SEQ ID NO: 1)

c) Rd designa *H. influenzae* cepa Rd.

Se pueden producir fácilmente diferentes fragmentos de pE en *E. coli*

Ocho secuencias de ADNc derivadas de la pE de longitud completa se clonaron en pET26b(+) y se expresaron en *E. coli*. Las proteínas resultantes se purificaron en resinas de níquel con afinidad por las etiquetas de histidina (Figura 8). Las proteínas recombinantes cubrieron todo el producto de proteína pE madura y sus longitudes y posiciones individuales fueron como se demostró en la Figura 8A y los productos purificados se divulgan en la Figura 8B.

pE es un factor de virulencia crucial en la otitis media aguda de rata (OMA)

Para investigar el papel de pE como factor de virulencia para NTHi, se desafió a las ratas en el oído medio con 10^5 a 10^9 NTHi 3655 Δ pe o la NTHi 3655 de tipo silvestre correspondiente (Figura 9). Curiosamente, se requirió de 100 a 1.000 veces más de NTHi 3655 Δ pe para inducir una OMA similar en comparación con la bacteria de tipo silvestre. Por lo tanto, pE es un factor de virulencia crucial para la OMA inducida por NTHi.

pE es altamente inmunogénica en una población definida

Para medir los niveles de anticuerpos en niños y donantes de sangre, se purificó pE(A) recombinante de *E. coli* (Figura 5B) y se usó en un ELISA. Los resultados de los análisis de ELISA de anticuerpos contra pE(A) se muestran en la Figura 9. Los niños menores de 6 meses de edad tenían IgG e IgA detectables contra pE(A). Los anticuerpos IgG mostraron niveles máximos en niños de 5 a 10 años de edad. Por el contrario, los anticuerpos IgA aumentaron gradualmente con el aumento de la edad y los valores más altos se detectaron en el grupo de edad de 70 a 80 años.

Epítomos útiles

Los epítomos de células B de una proteína se localizan principalmente en su superficie. Para predecir los epítomos de células B de los polipéptidos de la proteína E, se combinaron dos métodos: predicción de estructura 2D y predicción del índice antigénico. La predicción de la estructura 2D se realizó utilizando el programa PSIPRED (David Jones, Grupo de Bioinformática de Brunel, Departamento de Ciencias Biológicas, Universidad de Brunel, Uxbridge UB8 3PH, Reino Unido). El índice antigénico se calculó sobre la base del método descrito por Jameson y Wolf (CABIOS 4: 181-186 [1988]). Los parámetros utilizados en este programa son el índice antigénico y la longitud mínima para un péptido antigénico. Se usó un índice antigénico de 0,9 para un mínimo de 5 aminoácidos consecutivos como umbral en el programa. Los péptidos que comprenden buenos y potenciales epítomos de células B se enumeran en la tabla 3. Estos pueden ser útiles (preferiblemente conjugados o unidos de forma recombinante a una proteína más grande) en una composición de la vacuna para la prevención de infecciones por NTHi, al igual que péptidos similares que comprenden mutaciones conservadoras (preferiblemente 70, 80, 85, 95, 99 o 100% idénticas a las secuencias de la tabla 3) o truncadas que comprenden 5 o más (por ejemplo, 6, 7, 8, 9, 10, 11, 12, 15, 20 o 25) aminoácidos de las mismas o extensiones que comprenden por ejemplo, 1, 2, 3, 5, 10 aminoácidos adicionales en los extremos terminales N y/o C del péptido del contexto nativo del polipéptido de la proteína E (SEQ ID NO: 1 o sus homólogos naturales como se muestra en la Tabla 2 anterior) que preservan un epítomo efectivo que puede provocar una respuesta inmune en un huésped contra el polipéptido de la proteína E.

Tabla 3: Epítomos potenciales de células B de la SEQ ID NO: 1 (o sus homólogos naturales)

Posición	Secuencia
21	QKAKQND (SEQ ID NO: 21)
21	QKVKQND (SEQ ID NO: 22)
21	QKAQQND (SEQ ID NO: 23)
21	QKVQQND (SEQ ID NO: 24)
59	DNQEPQ (SEQ ID NO: 25)
82	PEPKRYARSVRQ (SEQ ID NO: 26)
106	QIRTDYDEFWGG (SEQ ID NO: 27)
106	QVRTDYDEFWGG (SEQ ID NO: 28)
123	APKKQKKH (SEQ ID NO: 29)
136	PDTTL (SEQ ID NO: 30)

20

Conclusiones

La proteína pE de la membrana externa de *Haemophilus* expuesta en la superficie es un factor de virulencia crucial para la OMA inducida por NTHi, es altamente inmunogénica en una población definida y, por lo tanto, es una vacuna candidata muy adecuada para una variedad de enfermedades humanas.

25

Referencias

- Ruan, M., M. Akkoyunlu, A. Grubb, y A. Forsgren. 1990. Protein D of *Haemophilus influenzae*. A novel bacterial surface protein with affinity for human IgD. J. Immunol. 145: 3379.
- Janson, H., L.-O. Hedén, A. Grubb, M. Ruan, y A. Forsgren. 1991. Protein D, an immunoglobulin D-binding protein of *Haemophilus influenzae*: cloning, nucleotide sequence, and expression in *Escherichia coli*. Infect. Immun. 59: 119.
- Janson, H., Å. Melhus, A. Hermansson, y A. Forsgren. 1994. Protein D, the glycerophosphodiester phosphodiesterase from *Haemophilus influenzae* with affinity for human immunoglobulin D, influences virulence in a rat otitis model. Infect. Immun. 62: 4848.
- Janson, H., B. Carlén, A. Cervin, A. Forsgren, A. Björk-Magnusdóttir, S. Lindberg, y T. Runer. 1999. Effects on the ciliated epithelium of protein D-producing and -nonproducing nontypeable *Haemophilus influenzae* in nasopharyngeal tissue cultures. J. Infect. Dis. 180: 737.
- Ahren, I. L., H. Janson, A. Forsgren, K. Riesbeck. 2001. Protein D expression promotes the adherence and internalization of non-typeable *Haemophilus influenzae* into human monocytic cells. Microb. Pathog. 31: 151.
- Forsgren, A. y Grubb, A. (1979) Many bacterial species bind human IgD. J. Immunol. 122, 1468-1472 .
- Banck, G. y Forsgren, A. (1978) Many bacterial species are mitogenic for human blood lymphocytes. Scand. J. Immunol. 8, 347-354 .
- Calvert, J.E. y Calogeres, A. (1986) Characteristics of human B cells responsive to the T-independent mitogen *Branhamella catarrhalis*. Immunology 58, 37-41.
- Forsgren, A., Penta, A., Schlossman, S.F. y Tedder, T.F. (1988) *Branhamella catarrhalis* activates human B lymphocytes following interactions with surface IgD and class I major histocompatibility complex antigens. Cell. Immunol. 112, 78-88 .
- Sasaki, K. y Munson Jr., R.S. (1993) Protein D of *Haemophilus influenzae* is not a universal immunoglobulin D-binding protein. Infect. Immun. 61, 3026-3031.
- Forsgren, A., M. Brant, A. Möllenkvist, A. Muyombwe, H. Janson, N. Woin, y K. Riesbeck., 2001. Isolation and characterization of a novel IgD-binding protein from *Moraxella catarrhalis*. J. Immunol. 167: 2112.
- Gorg, A., C. Obermaier, G. Boguth, W. Weiss. 1999. Recent developments in two-dimensional gel electrophoresis with immobilized pH gradients: wide pH gradients up to pH 12, longer separation distances and simplified procedures. Electrophoresis 20: 712.

55

ES 2 809 167 T3

13. Berns, K. I. y C. A. Thomas, Jr. 1965. Isolation og high molecular weight DNA from *Haemophilus influenzae*. J. Mol. Biol. 11: 476.
14. Clark-Curtiss, J. E. , W. R. Jacobs, M. A. Docherty, L. R. Ritchie, y R. Curtiss, 3a. 1985. Molecular analysis of DNA and construction of genomic libraries of Mycobacterium leprae. J Bacteriol. 161: 1093 .
- 5 15. Fleischmann, R. D., M. D. Adams, O. White, R. A. Clayton, E. F. Kirkness, A. R. Kerlavage, C. J. Bult, J. F. Tomb, B. A. Dougherty, J. M. Merrick, et al., 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science 269: 496.
16. Nielsen, H., J. Engelbrecht, S. Brunak, y G. von Heijne. 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Engineering 10: 1.
- 10 17. Kyte, J. y R. F. Doolittle. 1982. A simple method for displaying the hydrophatic character of a protein. J. Mol. Biol. 157: 105.
18. Poje, G., y J. R. Redfield. 2003. Transformation of *Haemophilus influenzae*. Methods. Mol. Med. 71: 57-70 .
19. Hussain, M., K. Becker, C. Von Eiff, J. Schrenzel, G. Peters, y M. Herrmann. 2001.J. Bacteriol. 183(23): 6778-6786.
- 15 Listado de secuencias
- <110> Arne Forsgren et al
- <120> Una nueva proteína de *Haemophilus influenzae* expuesta en la superficie (proteína E; pE)
- 20 <130> 21030217
- <160> 11
- 25 <170> PatentIn versión 3.3
- <210> 1
- <211> 160
- <212> PRT
- 30 <213> *Haemophilus influenzae* (pE)
- <400> 1

ES 2 809 167 T3

Met Lys Lys Ile Ile Leu Thr Leu Ser Leu Gly Leu Leu Thr Ala Cys
 1 5 10 15

Ser Ala Gln Ile Gln Lys Ala Lys Gln Asn Asp Val Lys Leu Ala Pro
 20 25 30

Pro Thr Asp Val Arg Ser Gly Tyr Ile Arg Leu Val Lys Asn Val Asn
 35 40 45

Tyr Tyr Ile Asp Ser Glu Ser Ile Trp Val Asp Asn Gln Glu Pro Gln
 50 55 60

Ile Val His Phe Asp Ala Val Val Asn Leu Asp Arg Gly Leu Tyr Val
 65 70 75 80

Tyr Pro Glu Pro Lys Arg Tyr Ala Arg Ser Val Arg Gln Tyr Lys Ile
 85 90 95

Leu Asn Cys Ala Asn Tyr His Leu Thr Gln Ile Arg Thr Asp Phe Tyr
 100 105 110

Asp Glu Phe Trp Gly Gln Gly Leu Arg Ala Ala Pro Lys Lys Gln Lys
 115 120 125

Lys His Thr Leu Ser Leu Thr Pro Asp Thr Thr Leu Tyr Asn Ala Ala
 130 135 140

Gln Ile Ile Cys Ala Asn Tyr Gly Glu Ala Phe Ser Val Asp Lys Lys
 145 150 155 160

<210> 2

5 <211> 139

<212> PRT

<213> Fragmento de proteína de *Haemophilus influenzae* o proteína sintética (pE(A))

<400> 2

10

Lys Ala Lys Gln Asn Asp Val Lys Leu Ala Pro Pro Thr Asp Val Arg
 1 5 10 15

ES 2 809 167 T3

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

5 <210> 3
 <211> 28
 <212> PRT
 <213> Fragmento de proteína de *Haemophilus influenzae* o péptido sintético (pE41-68)

10 <400> 3

Ile Arg Leu Val Lys Asn Val Asn Tyr Tyr Ile Asp Ser Glu Ser Ile
 1 5 10 15
 Trp Val Asp Asn Gln Glu Pro Gln Ile Val His Phe
 20 25

15 <210> 4
 <211> 39
 <212> PRT
 <213> Fragmento de proteína de *Haemophilus influenzae* o péptido sintético pE22-60)

<400> 4

Lys Ala Lys Gln Asn Asp Val Lys Leu Ala Pro Pro Thr Asp Val Arg
 1 5 10 15
 Ser Gly Tyr Ile Arg Leu Val Lys Asn Val Asn Tyr Tyr Ile Asp Ser
 20 25 30

20 Glu Ser Ile Trp Val Asp Asn
 35

25 <210> 5
 <211> 74
 <212> PRT
 <213> Fragmento de proteína de *Haemophilus influenzae* o péptido sintético (pE22-95)

<400> 5

ES 2 809 167 T3

Lys Ala Lys Gln Asn Asp Val Lys Leu Ala Pro Pro Thr Asp Val Arg
 1 5 10 15
 Ser Gly Tyr Ile Arg Leu Val Lys Asn Val Asn Tyr Tyr Ile Asp Ser
 20 25 30
 Glu Ser Ile Trp Val Asp Asn Gln Glu Pro Gln Ile Val His Phe Asp
 35 40 45
 Ala Val Val Asn Leu Asp Arg Gly Leu Tyr Val Tyr Pro Glu Pro Lys
 50 55 60
 Arg Tyr Ala Arg Ser Val Arg Gln Tyr Lys
 65 70

<210> 6

<211> 104

5 <212> PRT

<213> Fragmento de proteína de *Haemophilus influenzae* o péptido sintético (pE22-125)

<400> 6

Lys Ala Lys Gln Asn Asp Val Lys Leu Ala Pro Pro Thr Asp Val Arg
 1 5 10 15
 Ser Gly Tyr Ile Arg Leu Val Lys Asn Val Asn Tyr Tyr Ile Asp Ser
 20 25 30
 Glu Ser Ile Trp Val Asp Asn Gln Glu Pro Gln Ile Val His Phe Asp
 35 40 45
 Ala Val Val Asn Leu Asp Arg Gly Leu Tyr Val Tyr Pro Glu Pro Lys
 50 55 60
 Arg Tyr Ala Arg Ser Val Arg Gln Tyr Lys Ile Leu Asn Cys Ala Asn
 65 70 75 80
 Tyr His Leu Thr Gln Ile Arg Thr Asp Phe Tyr Asp Glu Phe Trp Gly
 85 90 95
 Gln Gly Leu Arg Ala Ala Pro Lys
 100

10

<210> 7

<211> 70

<212> PRT

15 <213> Fragmento de proteína de *Haemophilus influenzae* o péptido sintético (pE56-125)

<400> 7

ES 2 809 167 T3

Ile Trp Val Asp Asn Gln Glu Pro Gln Ile Val His Phe Asp Ala Val
 1 5 10 15
 Val Asn Leu Asp Arg Gly Leu Tyr Val Tyr Pro Glu Pro Lys Arg Tyr
 20 25 30
 Ala Arg Ser Val Arg Gln Tyr Lys Ile Leu Asn Cys Ala Asn Tyr His
 35 40 45
 Leu Thr Gln Ile Arg Thr Asp Phe Tyr Asp Glu Phe Trp Gly Gln Gly
 50 55 60
 Leu Arg Ala Ala Pro Lys
 65 70

<210> 8

<211> 105

5 <212> PRT

<213> Fragmento de proteína de *Haemophilus influenzae* o péptido sintético (pE56-160)

<400> 8

Ile Trp Val Asp Asn Gln Glu Pro Gln Ile Val His Phe Asp Ala Val
 1 5 10 15
 Val Asn Leu Asp Arg Gly Leu Tyr Val Tyr Pro Glu Pro Lys Arg Tyr
 20 25 30
 Ala Arg Ser Val Arg Gln Tyr Lys Ile Leu Asn Cys Ala Asn Tyr His
 35 40 45
 Leu Thr Gln Ile Arg Thr Asp Phe Tyr Asp Glu Phe Trp Gly Gln Gly
 50 55 60
 Leu Arg Ala Ala Pro Lys Lys Gln Lys Lys His Thr Leu Ser Leu Thr
 65 70 75 80
 Pro Asp Thr Thr Leu Tyr Asn Ala Ala Gln Ile Ile Cys Ala Asn Tyr
 85 90 95
 Gly Glu Ala Phe Ser Val Asp Lys Lys
 100 105

10

<210> 9

<211> 75

<212> PRT

15 <213> Fragmento de proteína de *Haemophilus influenzae* o péptido sintético (pE86-160)

<400> 9

Arg Tyr Ala Arg Ser Val Arg Gln Tyr Lys Ile Leu Asn Cys Ala Asn
 1 5 10 15

ES 2 809 167 T3

Tyr His Leu Thr Gln Ile Arg Thr Asp Phe Tyr Asp Glu Phe Trp Gly
 20 25 30
 Gln Gly Leu Arg Ala Ala Pro Lys Lys Gln Lys Lys His Thr Leu Ser
 35 40 45
 Leu Thr Pro Asp Thr Thr Leu Tyr Asn Ala Ala Gln Ile Ile Cys Ala
 50 55 60
 Asn Tyr Gly Glu Ala Phe Ser Val Asp Lys Lys
 65 70 75

5 <210> 10
 <211> 46
 <212> PRT
 <213> Fragmento de proteína de *Haemophilus influenzae* o péptido sintético (pE115-160)

10 <400> 10
 Phe Trp Gly Gln Gly Leu Arg Ala Ala Pro Lys Lys Gln Lys Lys His
 1 5 10 15
 Thr Leu Ser Leu Thr Pro Asp Thr Thr Leu Tyr Asn Ala Ala Gln Ile
 20 25 30
 Ile Cys Ala Asn Tyr Gly Glu Ala Phe Ser Val Asp Lys Lys
 35 40 45

15 <210> 11
 <211> 483
 <212> ADN
 <213> *Haemophilus influenzae*

<400> 11
 atgaaaaaaa ttattttaac attatcactt gggttactta ctgcctgttc tgctcaaadc 60
 caaaaggcta aacaaaatga tgtgaagctg gcaccgccga ctgatgtacg aagcggatat 120
 atacgtttgg taaagaatgt gaattattac atcgatagtg aatcgatctg ggtggataac 180
 caagagccac aaattgtaca ttttgatgca gtgggtgaatt tagatagggg attgtatggt 240
 tatcctgagc ctaaacgtta tgcacgttct gttcgtcagt ataagatcct gaattgtgca 300
 aattatcatt taactcaaat acgaactgat ttctatgatg aattttgggg acagggtttg 360
 cgggcagcac ctaaaaagca aaagaaacat acgttaagtt taacacctga tacaacgctt 420
 tataatgctg ctcagattat ttgtgcgaac tatgggtgaag cattttcagt tgataaaaaa 480
 20 taa 483

REIVINDICACIONES

- 5 1. Una composición de la vacuna para uso en el tratamiento o prevención de otitis media, sinusitis, infecciones del tracto respiratorio inferior o enfermedad pulmonar obstructiva crónica, causadas por *Haemophilus influenzae*, y que comprende una proteína expuesta en la superficie, que puede detectarse en *Haemophilus influenzae*, que tiene una secuencia de aminoácidos como se divulga en la SEQ ID NO: 1, 2, 6, 7, 8 o 9.
- 10 2. Una composición de la vacuna para uso de acuerdo con la reivindicación 1 que comprende una proteína inmunogénica recombinante basada en la proteína de la reivindicación 1, en la que los aminoácidos en las posiciones 1 a 21 de la SEQ ID NO: 1 se han eliminado o reemplazado por uno o más aminoácidos .
- 15 3. Una composición de la vacuna para uso de acuerdo con la reivindicación 2, en la que la proteína tiene una secuencia de aminoácidos de acuerdo con la SEQ ID NO: 2.
- 20 4. Una composición de la vacuna para uso de acuerdo con una cualquiera de las reivindicaciones 1-3, que comprende al menos un dímero, trímero o multímero de la proteína o polipéptido de acuerdo con una cualquiera de las reivindicaciones 1-3.
- 25 5. Una composición de la vacuna para uso de acuerdo con las reivindicaciones 1-4, que comprende además uno o más adyuvantes, vehículos, excipientes, aglutinantes, portadores, conservantes, agentes tamponantes, agentes emulsionantes, agentes humectantes o compuestos facilitadores de la transfección farmacéuticamente aceptables.
- 30 6. Una composición de la vacuna para uso de acuerdo con una cualquiera de las reivindicaciones 1-5, que comprende al menos una vacuna adicional.
- 35 7. Una composición de la vacuna para uso de acuerdo con una cualquiera de las reivindicaciones 1-6, que comprende una porción inmunogénica de otra molécula, y opcionalmente la porción inmunogénica de otra molécula se elige del grupo que comprende proteína D de *H. influenzae*, MID de *Moraxella catarrhalis*, UspA1 o UspA2 de *Moraxella catarrhalis* y proteína de la membrana externa de cualquier patógeno del tracto respiratorio.
- 40 8. Una composición de la vacuna para uso de acuerdo con una cualquiera de las reivindicaciones 1-3, que comprende un producto de fusión, en el que una proteína o polipéptido de acuerdo con una cualquiera de las reivindicaciones 1-5 está covalentemente, o por cualquier otro medio, unida a un proteína, carbohidrato o matriz.
- 45 9. Una composición de la vacuna para uso en el tratamiento o prevención de otitis media, sinusitis, infecciones del tracto respiratorio inferior o enfermedad pulmonar obstructiva crónica, causadas por *Haemophilus influenzae*, y que comprende una proteína expuesta en la superficie, que puede detectarse en *Haemophilus influenzae*, que comprende un polipéptido aislado que comprende la secuencia de aminoácidos de la SEQ ID NO: 1, opcionalmente en la que el polipéptido carece de un péptido señal (aminoácidos 1-20) de la SEQ ID NO: 1, y opcionalmente en la que el polipéptido tiene una o más de las siguientes sustituciones: Val en lugar de Ala en la posición 23 de la SEQ ID NO: 1, Glu en lugar de Lys en la posición 24 de la SEQ ID NO: 1, Met en lugar de Val en la posición 28 de la SEQ ID NO: 1, Thr en lugar de Ala en la posición 31 de la SEQ ID NO: 1, Val en lugar de Ile en la posición 41 de la SEQ ID NO: 1, Ala en lugar de Val en la posición 47 de la SEQ ID NO: 1, Lys en lugar de Arg en la posición 76 de la SEQ ID NO: 1, Val en lugar de Ile en la posición 107 de la SEQ ID NO: 1, Lys en lugar de Gly en la posición 152 de la SEQ ID NO: 1, Val en lugar de Ala en la posición 154 de la SEQ ID NO: 1)
- 50 10. Una composición de la vacuna para uso en el tratamiento o prevención de otitis media, sinusitis, infecciones del tracto respiratorio inferior o enfermedad pulmonar obstructiva crónica, causadas por *Haemophilus influenzae*, que comprende una proteína expuesta en la superficie, que puede detectarse en *Haemophilus influenzae*, que comprende el polipéptido aislado de la reivindicación 9, cuyo polipéptido (opcionalmente acoplado a un vehículo) es capaz de elevar una respuesta inmune que reconoce el polipéptido de la SEQ ID NO: 1 (o el polipéptido de la reivindicación 9, respectivamente), o es capaz de unirse a IgD humana.
- 55 11. Una composición de la vacuna para uso de acuerdo con la reivindicación 9, en la que dicho polipéptido es parte de una proteína de fusión más grande.
- 60 12. Una composición de la vacuna para uso de acuerdo con una cualquiera de las reivindicaciones 9 a 11 que comprende una cantidad eficaz del polipéptido de una cualquiera de las reivindicaciones 9 a 11 y un excipiente farmacéuticamente aceptable.
- 65 13. La composición de la vacuna para uso de acuerdo con la reivindicación 12, en la que dicha composición comprende al menos otro antígeno de *Haemophilus influenzae*.

14. La composición de la vacuna para uso de acuerdo con una cualquiera de las reivindicaciones 1-13, formulada con UspA1 y/o UspA2 de *Moraxella catarrhalis* o, formulada con proteína D de *Haemophilus influenzae*.

5 15. Una composición de la vacuna para uso de acuerdo con una cualquiera de las reivindicaciones 1-3, en la que la *Haemophilus influenzae* está encapsulada o no es tipificable.

16. Una composición de la vacuna para uso de acuerdo con la reivindicación 15, para la profilaxis o el tratamiento de otitis media, sinusitis o infecciones del tracto respiratorio inferior en niños y adultos que padecen, por ejemplo, enfermedad pulmonar obstructiva crónica (EPOC).

10

Figura 1 .

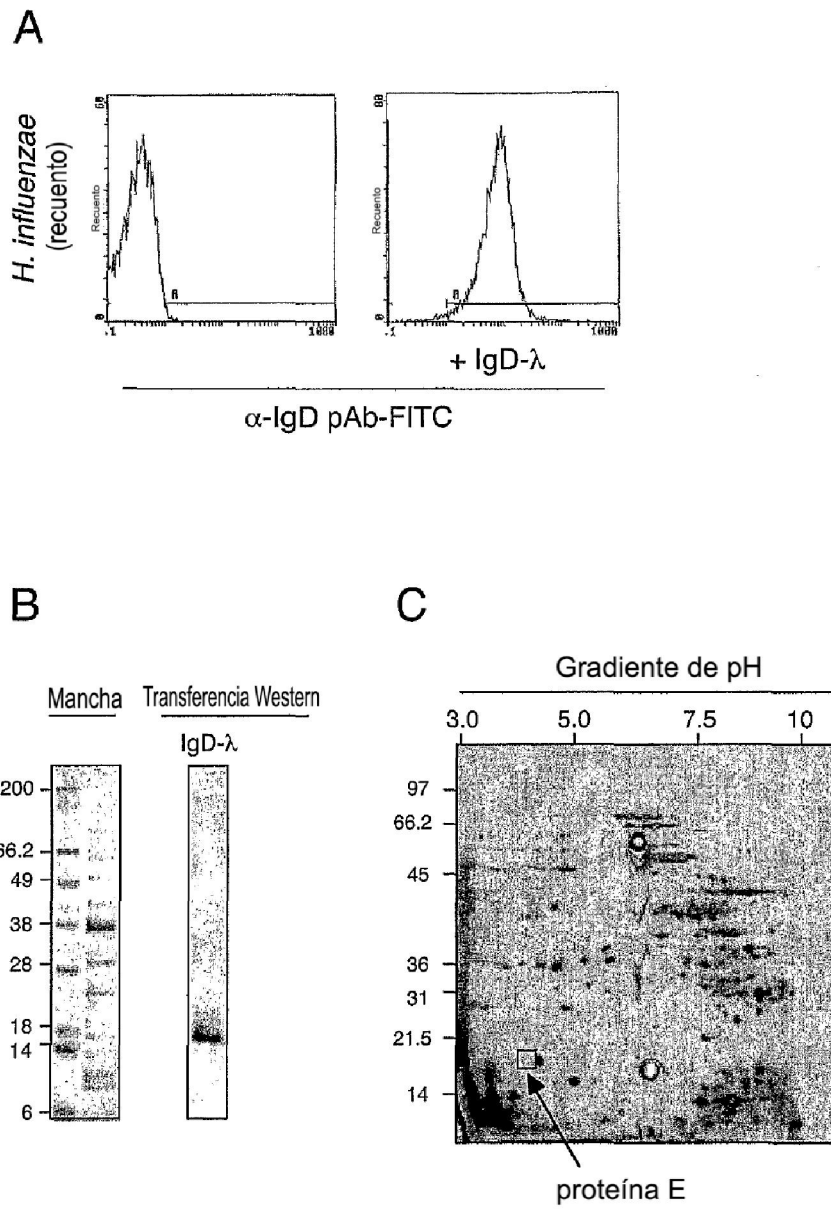


Figura 2.

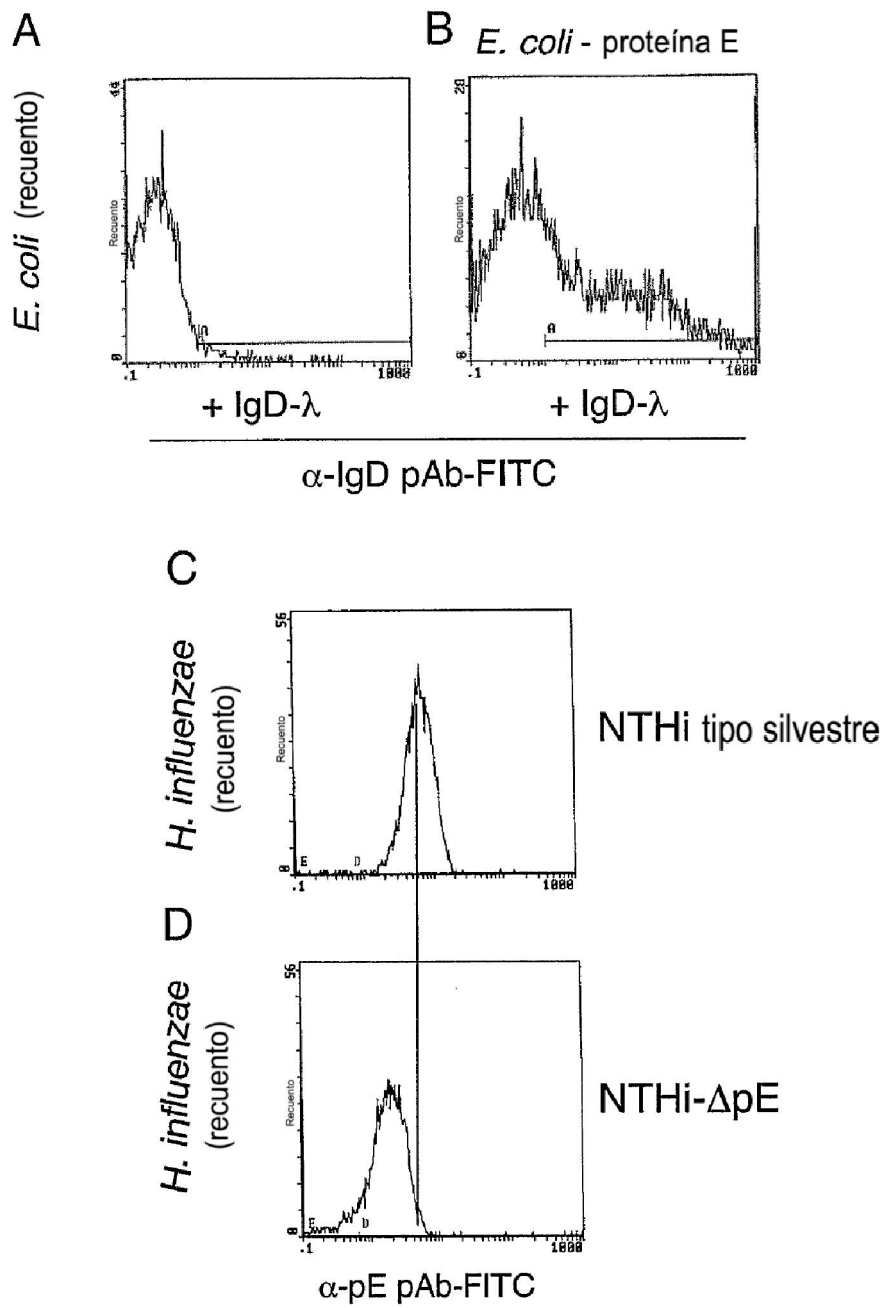


Figura 3.

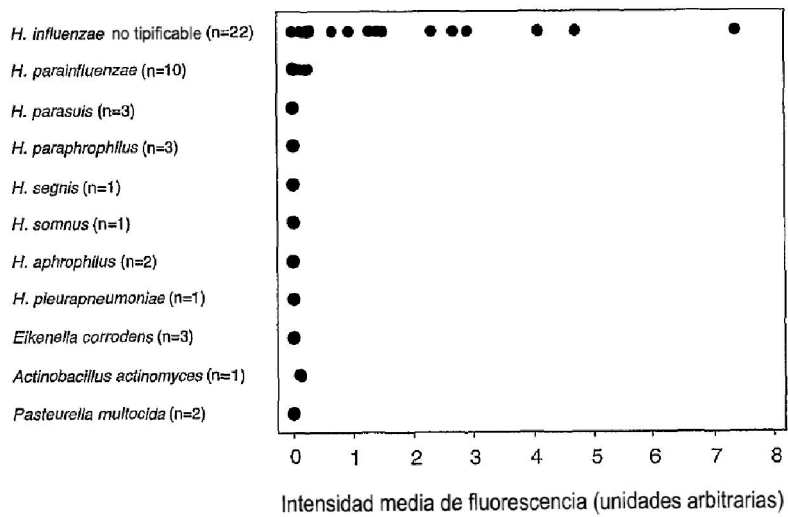


Figura 4 .

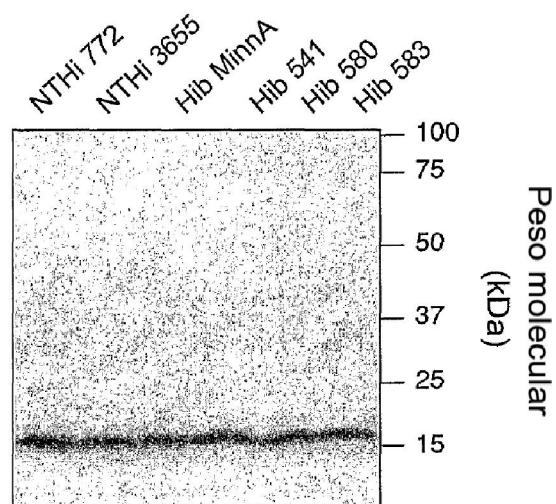


Figura 5.

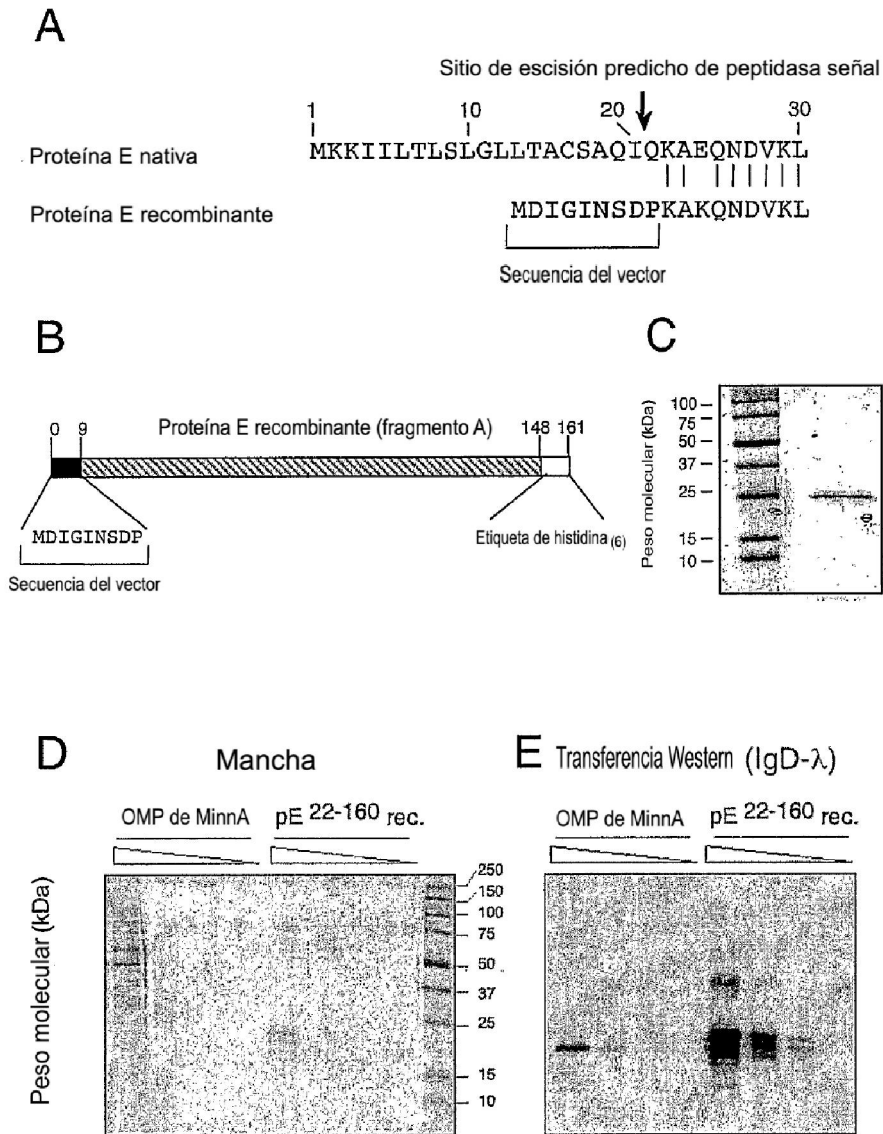


Figura 6.

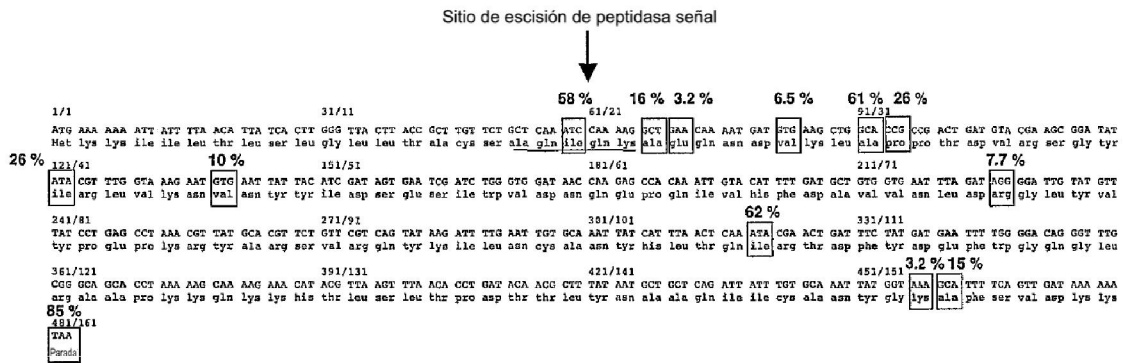


Figura 7.

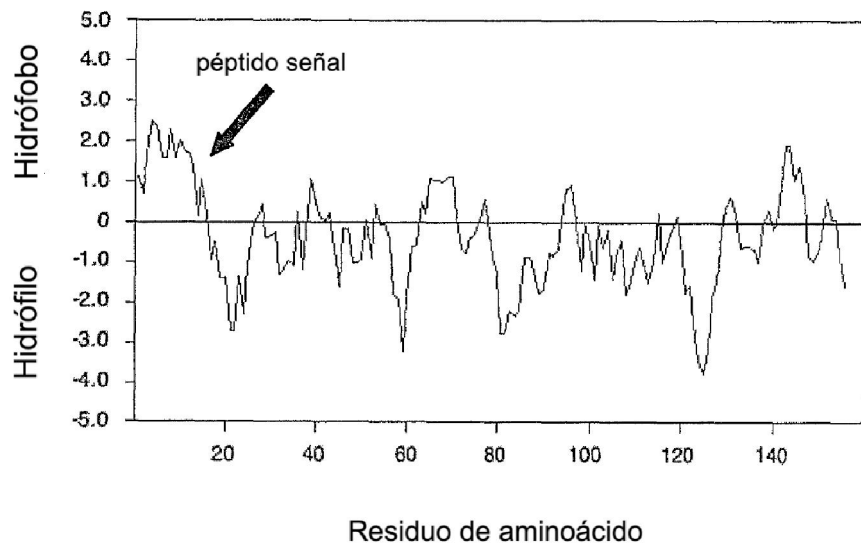
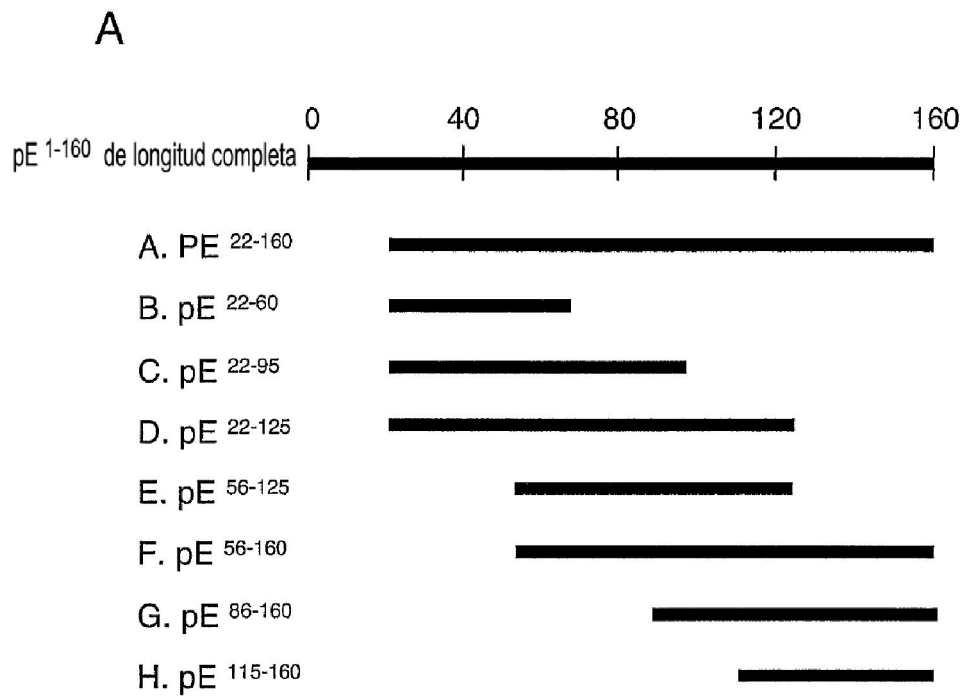


Figura 8.



B

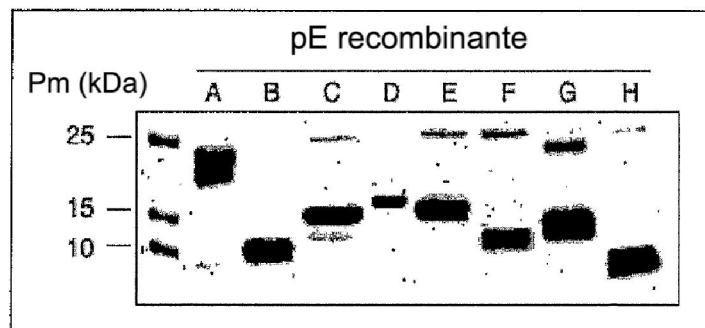


Figura 9.

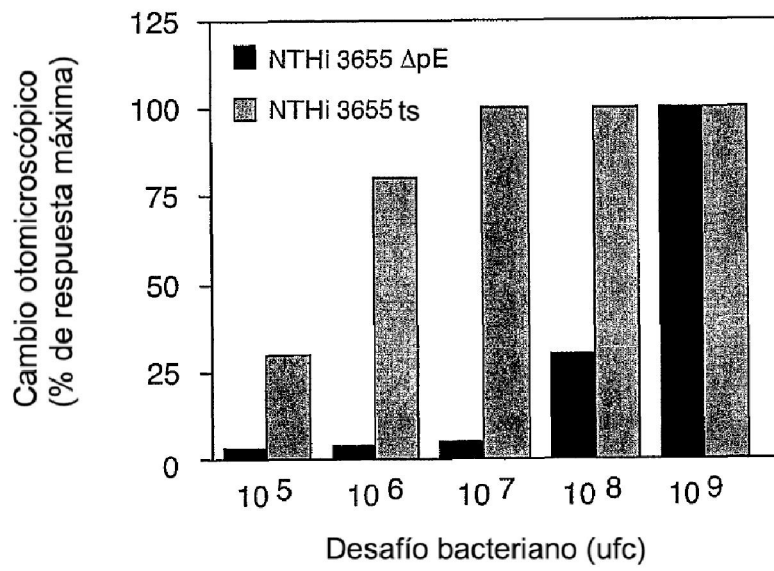


Figura 10.

