

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 454 266**

51 Int. Cl.:

**C12N 7/00** (2006.01)

**C12N 7/04** (2006.01)

**C07K 14/11** (2006.01)

**A61K 39/145** (2006.01)

12

## TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **20.05.2005 E 12172385 (2)**

97 Fecha y número de publicación de la concesión europea: **25.12.2013 EP 2522719**

54 Título: **Variantes de la hemaglutinina y la neuraminidasa de influenza**

30 Prioridad:

**25.05.2004 US 574553 P**

**28.02.2005 US 657554 P**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**10.04.2014**

73 Titular/es:

**MEDIMMUNE, LLC (50.0%)**

**One MedImmune Way**

**Gaithersburg, MD 20878, US y**

**THE GOVERNMENT OF THE UNITED STATES OF  
AMERICA NATIONAL INSTITUTES OF HEALTH  
(50.0%)**

72 Inventor/es:

**YANG, CHIN-FEN;**

**KEMBLE, GEORGE;**

**SUBBARAO, KANTA y**

**MURPHY, BRIAN**

74 Agente/Representante:

**LINAGE GONZÁLEZ, Rafael**

ES 2 454 266 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

Variantes de la hemaglutinina y la neuraminidasa de influenza

## 5 Antecedentes de la invención

Las vacunas contra diversas cepas en evolución de influenza son importantes desde un punto de vista de salud pública, así como comercialmente, puesto que cada año numerosos individuos se infectan con diferentes cepas y tipos de virus influenza. Los lactantes, ancianos y aquellos que tienen una atención sanitaria inadecuada y personas inmunocomprometidas corren un riesgo especial de muerte por tales infecciones. El problema de las infecciones por influenza se agrava porque las cepas de influenza novedosas evolucionan rápidamente y pueden propagarse entre diversas especies, necesitando de ese modo la producción continua de nuevas vacunas.

Se han producido numerosas vacunas que pueden producir una respuesta inmunitaria protectora específica para tales cepas de virus/virus influenza diferentes durante más de 50 años e incluyen vacunas de virus completo, vacunas de virus dividido, vacunas de antígeno de superficie y vacunas de virus atenuado vivo. Sin embargo, aunque formulaciones apropiadas de cualquiera de estos tipos de vacunas pueden producir una respuesta inmunitaria sistémica, las vacunas de virus atenuado vivo tienen la ventaja de también poder estimular la inmunidad mucosa local en el aparato respiratorio. Se ha realizado un considerable trabajo por los presentes inventores y colaboradores en la producción de virus influenza, y fragmentos de los mismos, para la producción de vacunas; véanse, por ejemplo, los documentos US 2004029251 y US 2005042229.

Debido a la continua aparición (o reaparición) de cepas de influenza diferentes, se desean continuamente nuevas vacunas contra influenza. Normalmente, tales vacunas se crean usando restos antigénicos de las cepas de virus recién aparecidas, por tanto, se desean altamente polipéptidos y polinucleótidos de cepas de virus novedosas, recién aparecidas o recién reaparecidas (especialmente secuencias de genes antigénicos).

El documento de Webby et al (2004/04/03), The Lancet, vol. 363, páginas 1099-1103, describe un uso de genética inversa para el desarrollo rápido de vacuna de influenza.

La presente descripción proporciona variantes de la hemaglutinina y la neuraminidasa de influenza nuevas y/o recién aisladas que son útiles en la producción de numerosos tipos de vacunas así como en investigación, diagnóstico, etc. Otros numerosos beneficios resultarán evidentes tras la revisión de lo siguiente.

## 35 Sumario de la invención

La invención en su sentido más amplio es tal como se define en las reivindicaciones independientes.

Se describe en el presente documento un polipéptido aislado o recombinante que se selecciona de: los polipéptidos codificados por una cualquiera de las secuencias de SEC ID N° 1 a SEC ID N° 10, uno cualquiera de los polipéptidos codificados por las SEC ID N° 1 a SEC ID N° 10; uno cualquiera de los polipéptidos de las SEC ID N° 11 a SEC ID N° 20; sólo los marcos de lectura abiertos de los polipéptidos de las SEC ID N° 11 de SEC ID N° 20; formas alternativas (por ejemplo, la forma madura sin el péptido señal, o sin las secuencias 5' y 3' fuera del marco de lectura abierto, o las secuencias tal como se expresan en la superficie de un virus (por ejemplo, influenza)) de los polipéptidos de SEC ID N° 11-20. Los polipéptidos descritos en el presente documento comprenden opcionalmente proteínas de fusión, proteínas con una secuencia líder, un polipéptido precursor, proteínas con una señal de secreción o una señal de localización, o proteínas con una etiqueta de epítipo, una etiqueta de E o una etiqueta de epítipo de His. Las secuencias descritas en el presente documento también se muestran en el apéndice 1 y en las listas de secuencias en el presente documento. Las secuencias de hemaglutinina descritas en el presente documento comprenden secuencias con sitios de escisión polibásicos modificados (que permiten de ese modo el crecimiento de los virus en huevos). Las secuencias de polipéptido de hemaglutinina de SEC ID N° 11-20 comprenden las secuencias de péptido señal amino terminal endógenas, sin embargo, las secuencias de polipéptido de hemaglutinina descritas en el presente documento también incluyen la forma madura (péptido señal amino terminal escindido) de los polipéptidos de hemaglutinina. Los sitios de escisión de cualquier secuencia de polipéptido de hemaglutinina de cualquier cepa de influenza pueden predecirse o medirse rutinariamente usando cualquiera de varios métodos en la técnica.

También se describe en el presente documento una composición con uno o más polipéptidos enumerados anteriormente, o fragmentos de los mismos. Se describen además en el presente documento polipéptidos a los que se une específicamente un antisuero policlonal generado contra al menos 1 antígeno que comprende al menos una secuencia de aminoácidos descrita anteriormente, o un fragmento de la misma. Tales anticuerpos específicos para los polipéptidos descritos anteriormente también se describen en el presente documento. Los polipéptidos descritos en el presente documento son opcionalmente inmunogénicos.

Se describen además composición inmunogénicas que comprenden una cantidad inmunológicamente eficaz de uno o más de cualquiera de los polipéptidos descritos anteriormente así como métodos para estimular el sistema

inmunitario de un individuo para producir una respuesta inmunitaria protectora contra virus influenza administrando al individuo una cantidad inmunológicamente eficaz de cualquiera de los polipéptidos anteriores en un vehículo fisiológicamente aceptable.

- 5 La invención incluye virus influenza recombinante tal como se caracteriza mediante las reivindicaciones adjuntas que comprende uno o más de los polipéptidos o polinucleótidos anteriores, además de composiciones inmunogénicas que comprenden una cantidad inmunológicamente eficaz de tal virus influenza recombinante. El virus influenza recombinante de la invención para su uso para estimular el sistema inmunitario de un individuo para producir una respuesta inmunitaria protectora contra virus influenza, a través de la administración de una cantidad  
10 inmunológicamente eficaz de tal virus influenza recombinante en un vehículo fisiológicamente aceptable también es parte de la invención.

También se describe en el presente documento una composición de materia que tiene dos o más ácidos nucleicos descritos anteriormente (por ejemplo, una biblioteca que comprende al menos aproximadamente 2, 5, 10, 50 o más  
15 ácidos nucleicos). Tales composiciones pueden producirse opcionalmente escindiendo uno o más ácido(s) nucleico(s) descrito(s) anteriormente, (por ejemplo, de manera mecánica, química, enzimática con una endonucleasa de restricción/ARNasa/ADNasa, etc.). Otras composiciones incluyen, por ejemplo, composiciones producidas incubando uno o más de los ácido(s) nucleico(s) descrito(s) anteriormente en presencia de desoxirribonucleótidos trifosfato y una ácido nucleico polimerasa termoestable.

20 También se describen en el presente documento células que comprenden al menos uno de los ácidos nucleicos descritos anteriormente, o un producto o fragmento amplificado o escindido de los mismos. Tales células pueden expresar opcionalmente un polipéptido codificado por tal ácido nucleico. Se describen además en el presente documento vectores (por ejemplo, plásmidos, cósmidos, fagos, virus, fragmentos de virus, etc.) que comprenden  
25 cualquiera de los ácidos nucleicos descritos anteriormente. Tales vectores pueden comprender opcionalmente un vector de expresión. Los vectores de expresión preferidos incluyen, pero no se limitan a, vectores que comprenden el promotor de pol I y secuencias terminadoras o vectores que usan los promotores tanto de pol I como de pol II "el sistema de promotores de pol I/pol II" (por ejemplo, Zobel *et al.*, Nucl. Acids Res. 1993,21:3607; documento US 20020164770; Neumann *et al.*, Proc. Natl. Acad. Sci. USA 1999, 96:9345; Fodor *et al.*, J. Virol. 1999, 73:9679; y  
30 documento US 20030035814). También se describen células transducidas por tales vectores.

En algunas realizaciones, la invención abarca un virus (por ejemplo, un virus influenza) que comprende uno o más ácidos nucleicos descritos anteriormente (por ejemplo, que codifican la hemaglutinina y/o la neuraminidasa) tal como se caracteriza mediante las reivindicaciones adjuntas. Composiciones inmunogénicas que comprenden tal virus  
35 también son parte de la presente invención tal como se caracteriza mediante las reivindicaciones adjuntas. Un virus de este tipo es un virus con reagrupamiento 6:2 (por ejemplo, que comprende 6 regiones codificantes de genes que codifican segmentos de genoma internos de A/AA/6/60 y 2 regiones codificantes que codifican una hemaglutinina que comprende la SEC ID N° 11 y una neuraminidasa que comprende la SEC ID N° 12). Un virus influenza reagrupado de la invención incluye un virus influenza con reagrupamiento 6:2, en el que dicho virus comprende 6  
40 regiones codificantes de genes que codifican segmentos de genoma internos de A/Ann Arbor/6/60 y 2 regiones codificantes de genes, en el que las 2 regiones codificantes de genes son HA que comprende la SEC ID N° 11 y NA que comprende la SEC ID N° 12. Métodos tal como se caracterizan mediante las reivindicaciones adjuntas de producción de virus influenza recombinante a través del cultivo de una célula huésped que alberga un virus influenza en un medio de cultivo adecuado en condiciones que permiten la expresión del ácido nucleico y el aislamiento del  
45 virus influenza recombinante de una o más de la(s) célula(s) huésped o el medio también son parte de la invención.

En otras realizaciones en el presente documento, la invención comprende composiciones inmunogénicas que tienen una cantidad inmunológicamente eficaz de cualquiera de los virus influenza recombinantes descritos anteriormente tal como se caracterizan mediante las reivindicaciones adjuntas. Otras realizaciones incluyen el virus influenza reagrupado de la invención para su uso en métodos para estimular el sistema inmunitario de un individuo para producir una respuesta inmunitaria protectora contra virus influenza administrando al individuo una cantidad  
50 inmunológicamente eficaz de cualquiera de los virus influenza recombinantes descritos anteriormente (opcionalmente en un vehículo fisiológicamente eficaz).

55 También se describen en el presente documento métodos de producción de un polipéptido aislado o recombinante cultivando cualquier célula huésped anterior, en un medio de cultivo adecuado en condiciones que permiten la expresión de ácido nucleico, y aislando el polipéptido de una o más de las células huésped o el medio en el que se hacen crecer las células.

60 Composiciones inmunogénicas tal como se caracterizan mediante las reivindicaciones adjuntas también son características de la invención. Por ejemplo, composiciones inmunogénicas que comprenden uno cualquiera o más virus descritos anteriormente tal como se caracterizan mediante las reivindicaciones adjuntas (por ejemplo, junto con uno o más componente(s) de administración farmacéuticamente aceptables.

65 El virus influenza reagrupado de la invención para su uso en un método de tratamiento profiláctico de una infección viral (por ejemplo, gripe viral) en un sujeto a través de la administración de uno cualquiera o más virus descritos

anteriormente o composiciones inmunogénicas tal como se mencionan en las reivindicaciones en una cantidad eficaz para producir una respuesta inmunogénica contra la infección viral también es parte de la presente invención. Los sujetos para tal tratamiento pueden incluir mamíferos (por ejemplo, seres humanos). Tal virus reagrupado de la invención para su uso en tales métodos también puede ser para la administración *in vivo* al sujeto así como para la

- 5 administración *in vitro* o *ex vivo* a una o más célula(s) del sujeto. Adicionalmente, tal virus reagrupado de la invención para su uso en tales métodos también puede comprender una composición del virus y un excipiente farmacéuticamente aceptable para su administración al sujeto en una cantidad eficaz para tratar profilácticamente la infección viral.
- 10 También se describen en el presente documento composiciones de materia que comprenden secuencias de ácido nucleico que codifican para polipéptidos de hemaglutinina y/o neuraminidasa de una o más cepa(s) de influenza pandémica(s) y secuencias de ácido nucleico que codifican para uno o más polipéptido(s) de A/Ann Arbor/6/60. Se describen además composiciones de materia que comprenden secuencias de ácido nucleico que codifican para
- 15 polipéptidos de hemaglutinina y/o neuraminidasa de una o más cepa(s) de influenza pandémica(s) y secuencias de ácido nucleico que codifican para uno o más polipéptido(s) de PR8 o A/Ann Arbor/6/60. Tales secuencias pueden incluir las enumeradas en la lista de secuencias en el presente documento. Adicionalmente, estas composiciones incluyen composiciones de materia que comprende secuencias que codifican para hemaglutinina y/o neuraminidasa de una o más cepa(s) de influenza pandémica(s) y secuencias de ácido nucleico que codifican para una cepa de
- 20 estructura principal seleccionada en un reagrupamiento 6:2. Tales composiciones descritas en el presente documento incluyen secuencias que codifican la hemaglutinina y la neuraminidasa seleccionadas del listado de secuencias en el presente documento y una cepa de estructura principal, siendo la cepa de estructura principal PR8 o A/Ann Arbor/6/60. También se describen composiciones tal como se describió anteriormente en las que la hemaglutinina comprende un sitio de escisión polibásico modificado. La invención también incluye una vacuna contra influenza atenuada viva que comprende tales composiciones anteriores tal como se caracterizan mediante las
- 25 reivindicaciones adjuntas.

Estos y otros objetos y características de la invención resultarán más completamente evidentes cuando se lea la siguiente descripción detallada conjuntamente con el apéndice y las figuras adjuntas.

### 30 Breve descripción de los dibujos

Figura 1: Muestra modificaciones realizadas mediante ingeniería genética en el gen de HA de VN/1203/2004 para eliminar el sitio de escisión polibásico.

- 35 Figura 2: Presenta resultados que muestran que virus reagrupados H5N1 *ca* administrados por vía intranasal no se replican en pollos.

Figura 3: Ilustra que los candidatos a vacuna H5N1/AA *ca* no son letales para ratones.

- 40 Figura 4: Ilustra que los virus reagrupados H5N1 *ca* 1997 y 2004 están restringidos en la replicación en ratones.

Figura 5: Ilustra que los virus influenza H5N1/AA *ca* reagrupados están restringidos en la replicación en pulmones de ratones.

- 45 Figura 6: Muestra los títulos de Ac HAI séricos producidos en ratones tras una única dosis i.n. de vacuna.

Figura 7: Muestra los títulos de Ac neutralizantes séricos producidos en ratones tras una única dosis i.n. de vacuna.

- 50 Figura 8: Ilustra que virus reagrupados H5N1 *ca* protegen a los ratones frente a exposiciones letales con 50, 500 o 5000 DL<sub>50</sub> de virus H5N1 de tipo natural.

Figura 9: Ilustra la eficacia de la protección frente a la replicación pulmonar de los virus de exposición H5N1 homólogos y heterólogos en ratones.

- 55 Figura 10: Ilustra la eficacia de la protección frente a la replicación de los virus de exposición H5N1 homólogos y heterólogos en el aparato respiratorio superior de ratones.

Figura 11: Ilustra la eficacia de protección conferida por la vacuna H5N1 *ca* 2004 frente a la exposición a alta dosis (10<sup>5</sup>DICT<sub>50</sub>) con virus wt H5N1 homólogos o heterólogos en ratones.

- 60 Figura 12: Ilustra la eficacia de la protección conferida por las vacunas H5N1 *ca* 1997 y 2003 frente a exposiciones a dosis altas (10<sup>5</sup>DICT<sub>50</sub>) con virus H5N1 de tipo natural homólogos o heterólogos en ratones.

- 65 Figura 13: Ilustra la eficacia de la protección conferida por la vacuna H5N1 *ca* 2004 frente a dosis bajas o altas de exposiciones al virus H5N1 de tipo natural homólogo en ratones.

## Descripción detallada

### Definiciones

- 5 A menos que se defina lo contrario, todos los términos técnicos y científicos usados en el presente documento tienen el mismo significado que entiende comúnmente un experto habitual en la técnica a la que se refiere la invención. Las siguientes definiciones complementan aquéllas en la técnica y se refieren a la presente solicitud y no deben atribuirse necesariamente a ningún caso relacionado o no relacionado, por ejemplo, a ninguna solicitud o patente de propiedad conjunta. Aunque puede usarse cualquier método y material similar o equivalente a los descritos en el
- 10 presente documento en la puesta en práctica para someter a prueba la presente invención, se describen en el presente documento los materiales y métodos preferidos. Por consiguiente, la terminología usada en el presente documento es para el fin de describir realizaciones particulares sólo, y no pretende ser limitante.
- 15 Tal como se usa en esta memoria descriptiva y las reivindicaciones adjuntas, las formas singulares “un”, “una” y “el/la” incluyen referentes plurales a menos que el contexto dicte claramente lo contrario. Por tanto, por ejemplo, la referencia a “un virus” incluye una pluralidad de virus; la referencia a una “célula huésped” incluye mezclas de células huésped y similares.
- 20 Una “secuencia de aminoácidos” es un polímero de residuos de aminoácido (una proteína, polipéptido, etc.) o una cadena de caracteres que representa un polímero de aminoácidos, dependiendo del contexto.
- Los términos y expresiones “ácido nucleico”, “polinucleótido”, “secuencia de polinucleótidos” y “secuencia de ácido nucleico” se refieren a polímeros de desoxirribonucleótidos o ribonucleótidos monocatenarios o bicatenarios, quimeras o análogos de los mismos, o una cadena de caracteres que representan tales, dependiendo del contexto.
- 25 Tal como se usa en el presente documento, el término incluye opcionalmente polímeros de análogos de nucleótidos que se producen de manera natural que tienen la naturaleza esencial de nucleótidos naturales porque hibridan con ácidos nucleicos monocatenarios de una manera similar a nucleótidos que se producen de manera natural (por ejemplo, ácidos nucleicos peptídicos). A menos que se indique lo contrario, una secuencia de ácido nucleico particular abarca opcionalmente secuencias complementarias además de la secuencia explícitamente indicada. A
- 30 partir de cualquier secuencia de polinucleótido especificada, puede determinarse o bien el ácido nucleico dado o bien la secuencia de polinucleótido complementaria (por ejemplo, el ácido nucleico complementario).
- La expresión “ácido nucleico” o “polinucleótido” también abarca cualquier cadena física de unidades de monómero que puede hacerse corresponder con una cadena de nucleótidos, incluyendo un polímero de nucleótidos (por ejemplo, un polímero de ADN o ARN típico), PNA, oligonucleótidos modificados (por ejemplo, oligonucleótidos que comprenden bases que no son típicas para ADN o ARN biológico en disolución, tales como oligonucleótidos 2'-O-metilados) y similares. Un ácido nucleico puede ser, por ejemplo, monocatenario o bicatenario.
- 35 Una “subsecuencia” es cualquier parte de una secuencia entera, hasta e incluyendo la secuencia completa. Normalmente, una subsecuencia comprende menos de la secuencia de longitud completa. Una “subsecuencia única” es una subsecuencia que no se encuentra en ninguna secuencia de polinucleótido o polipéptido de influencia determinada previamente.
- 40 El término “variante” con respecto a un polipéptido se refiere a una secuencia de aminoácidos que está alterada en uno o más aminoácidos con respecto a una secuencia de referencia. La variante puede tener cambios “conservativos”, cuando un aminoácido sustituido tiene propiedades químicas o estructurales similares, por ejemplo, la sustitución de leucina por isoleucina. Alternativamente, una variante puede tener cambios “no conservativos”, por ejemplo, la sustitución de una glicina por un triptófano. La variación menor análoga también puede incluir delección o inserción de aminoácidos, o ambos. Puede encontrarse orientación para determinar qué residuos de aminoácido
- 45 pueden sustituirse, insertarse o deleccionarse sin eliminar la actividad biológica o inmunológica usando programas informáticos bien conocidos en la técnica, por ejemplo, el software DNASTAR. También se describen ejemplos de sustituciones conservadoras en el presente documento.
- 50 El término “gen” se usa de manera amplia para referirse a cualquier ácido nucleico asociado con una función biológica. Por tanto, los genes incluyen secuencias codificantes y/o secuencias reguladoras requeridas para su expresión. El término “gen” se aplica a una secuencia genómica específica, así como a un ADNc o un ARNm codificado por esa secuencia genómica.
- 55 Los genes también incluyen segmentos de ácido nucleico no expresados que, por ejemplo, forman secuencias de reconocimiento para otras proteínas. Las secuencias reguladoras no expresadas incluyen “promotores” y “potenciadores”, a los que se unen proteínas reguladoras tales como factores de transcripción, dando como resultado la transcripción de secuencias adyacentes o cercanas. Un promotor o potenciador “específico de tejido” es uno que regula la transcripción en un tipo de célula o tipo de tejido específico, o tipos.
- 60 “Expresión de un gen” o “expresión de un ácido nucleico” significa la transcripción de ADN en ARN (incluyendo opcionalmente la modificación del ARN, por ejemplo, corte y empalme alternativo), la traducción del ARN en un
- 65

polipéptido (incluyendo posiblemente la modificación posterior del polipéptido, por ejemplo, modificación postraducciona), o tanto la transcripción como la traducción, según indique el contexto.

Un “marco de lectura abierto” u “ORF” es un posible marco de lectura traduccional de ADN o ARN (por ejemplo, de un gen), que puede traducirse en un polipéptido. Es decir, el marco de lectura no está interrumpido por codones de terminación. Sin embargo, debe indicarse que el término ORF no indica necesariamente que el polinucleótido se traduzca de hecho en un polipéptido.

El término “vector” se refiere a los medios mediante los cuales puede propagarse y/o transferirse un ácido nucleico entre organismos, células o componentes celulares. Los vectores incluyen plásmidos, virus, bacteriófagos, provirus, fagémidos, transposones, cromosomas artificiales y similares, que se replican de manera autónoma o pueden integrarse en un cromosoma de una célula huésped. Un vector también puede ser un polinucleótido de ARN desnudo, un polinucleótido de ADN desnudo, un polinucleótido compuesto por tanto ADN como ARN dentro de la misma hebra, un ADN o ARN conjugado con polilisina, un ADN o ARN conjugado con péptidos, un ADN conjugado con liposomas o similares, que no se replica de manera autónoma. En muchos, pero no todos los casos comunes, los vectores son plásmidos.

Un “vector de expresión” es un vector, tal como un plásmido que puede promover la expresión, así como la replicación de un ácido nucleico incorporado en el mismo. Normalmente, el ácido nucleico que va a expresarse está “operativamente unido” a un promotor y/o potenciador, y se ve sometido a control regulatorio de la transcripción por el promotor y/o potenciador.

Un “vector de expresión bidireccional” se caracteriza por dos promotores alternativos orientados en la dirección opuesta en relación con un ácido nucleico situado entre los dos promotores, de manera que puede iniciarse la expresión en ambas orientaciones dando como resultado, por ejemplo, la transcripción de tanto ARN de hebra positiva (+) o sentido, como de hebra negativa (-) o antisentido.

Un “polipéptido” es un polímero que comprende dos o más residuos de aminoácido (por ejemplo, un péptido o una proteína). El polímero puede comprender opcionalmente modificaciones tales como glicosilación o similares. Los residuos de aminoácido del polipéptido pueden ser naturales o no naturales y pueden estar no sustituidos, no modificados, sustituidos o modificados.

En el contexto de la invención, el término “aislado” se refiere a un material biológico, tal como un virus, un ácido nucleico o una proteína, que está sustancialmente libre de componentes que normalmente lo acompañan o interaccionan con el mismo en su entorno que se produce de manera natural. El material biológico aislado comprende opcionalmente material adicional no encontrado con el material biológico en su entorno natural, por ejemplo, un virus de tipo natural o una célula. Por ejemplo, si el material está en su entorno natural, tal como una célula, el material puede colocarse en una ubicación en la célula (por ejemplo, genoma o elemento genético) no nativa para tal material encontrado en ese entorno. Por ejemplo, un ácido nucleico que se produce de manera natural (por ejemplo, una secuencia codificante, un promotor, un potenciador, etc.) se aísla si se introduce por medios que no se producen de manera natural en un locus del genoma (por ejemplo, un vector, tal como un vector de virus o plásmido, o amplicón) no nativo para ese ácido nucleico. Tales ácidos nucleicos se denominan también ácidos nucleicos “heterólogos”. Un virus aislado, por ejemplo, está en un entorno (por ejemplo, un sistema de cultivo celular, o purificado del cultivo celular) distinto del entorno nativo del virus de tipo natural (por ejemplo, la nasofaringe de un individuo infectado).

El término “quimérico” o “quimera”, cuando se refiere a un virus, indica que el virus incluye componentes genéticos y/o de polipéptido derivados de más de una fuente o cepa viral original. De manera similar, el término “quimérico” o “quimera”, cuando se refiere a una proteína viral, indica que la proteína incluye componentes de polipéptido (es decir, subsecuencias de aminoácidos) derivados de más de una fuente o cepa viral original.

El término “recombinante” indica que el material (por ejemplo, un ácido nucleico o una proteína) se ha alterado de manera artificial o sintética (de manera no natural) por la intervención humana. La alteración puede realizarse en el material dentro de, o extraído de, su estado o entorno natural. Específicamente, por ejemplo, un virus influenza es recombinante cuando se produce por la expresión de un ácido nucleico recombinante. Por ejemplo, un “ácido nucleico recombinante” es uno que se prepara recombinando ácidos nucleicos, por ejemplo durante la clonación, transposición del ADN u otros procedimientos, o por mutagénesis química u otra; un “polipéptido recombinante” o “proteína recombinante” es un polipéptido o proteína que se produce por la expresión de un ácido nucleico recombinante; y un “virus recombinante”, por ejemplo un virus influenza recombinante, se produce por la expresión de un ácido nucleico recombinante.

El término “reagrupado”, cuando se refiere a un virus, indica que el virus incluye componentes genéticos y/o de polipéptido derivados de más de una fuente o cepa viral original. Por ejemplo, un reagrupamiento 7:1 incluye 7 segmentos genómicos virales (o segmentos genéticos) derivados de un primer virus original, y un único segmento genómico viral complementario, por ejemplo, que codifica para una hemaglutinina o neuraminidasa. Un reagrupamiento 6:2 incluye 6 segmentos genómicos, lo más comúnmente los 6 genes internos de un primer virus

original, y dos segmentos complementarios, por ejemplo, hemaglutinina y neuraminidasa, de un virus original diferente.

El término “introducido” cuando se refiere a un ácido nucleico heterólogo o aislado, se refiere a la incorporación de un ácido nucleico en una célula eucariota o procariota en la que el ácido nucleico puede incorporarse en el genoma de la célula (por ejemplo, cromosoma, plásmido, plástido o ADN mitocondrial), convertirse en un replicón autónomo o expresarse de manera transitoria (por ejemplo ARNm transfectado). El término incluye métodos tales como “infección”, “transfección”, “transformación” y “transducción”. En el contexto de la invención, puede emplearse una variedad de métodos para introducir ácidos nucleicos en células, incluyendo electroporación, precipitación con fosfato de calcio, transfección mediada por lípidos (lipofección), etc.

El término “célula huésped” significa una célula que contiene un ácido nucleico heterólogo, tal como un vector, y soporta la replicación y/o expresión del ácido nucleico. Las células huésped pueden ser células procariotas tales como *E. coli*, o células eucariotas tales como células de levaduras, insectos, anfibios, aves o mamíferos, incluyendo células humanas. Las células huésped a modo de ejemplo pueden incluir, por ejemplo, células Vero (riñón de mono verde africano), células BHK (riñón de cría de hámster), células de riñón de pollo primarias (PCK), células de riñón canino Madin-Darby (MDCK), células de riñón bovino Madin-Darby (MDBK), células 293 (por ejemplo, células 293T) y células COS (por ejemplo, células COS1, COS7), etc.

Una “cantidad inmunológicamente eficaz” de virus influenza es una cantidad suficiente para potenciar la propia respuesta inmunitaria de un individuo (por ejemplo, de un ser humano) frente a una exposición posterior a virus influenza. Pueden monitorizarse los niveles de inmunidad inducida, por ejemplo, midiendo las cantidades de anticuerpos séricos y/o secretores neutralizantes, por ejemplo, mediante ensayo de neutralización de placas, de fijación del complemento, ensayo inmunoabsorbente ligado a enzimas o de microneutralización.

Una “respuesta inmunitaria protectora” frente a virus influenza se refiere a una respuesta inmunitaria presentada por un individuo (por ejemplo, un ser humano) que es protectora frente a la enfermedad cuando el individuo se expone posteriormente a y/o se infecta con tal virus influenza. En algunos casos, el virus influenza (por ejemplo, que circula de manera natural) puede provocar todavía infección, pero no puede provocar una infección grave. Normalmente, la respuesta inmunitaria protectora da como resultado niveles detectables de anticuerpos secretores y séricos engendrados por el huésped que pueden neutralizar el virus de la misma cepa y/o subgrupo (y posiblemente también de un subgrupo y/o cepa que no es de vacuna, diferente), *in vitro* e *in vivo*.

Tal como se usa en el presente documento, un “anticuerpo” es una proteína que comprende uno o más polipéptidos sustancial o parcialmente codificados por genes de inmunoglobulina o fragmentos de genes de inmunoglobulina. Los genes de inmunoglobulina reconocidos incluyen los genes de regiones constantes kappa, lambda, alfa, gamma, delta, epsilon y mu, así como miríadas de genes de regiones variables de inmunoglobulina. Las cadenas ligeras se clasifican como o bien kappa o bien lambda. Las cadenas pesadas se clasifican como gamma, mu, alfa, delta o epsilon, que a su vez definen las clases de inmunoglobulinas, IgG, IgM, IgA, IgD e IgE, respectivamente. Una unidad estructural de inmunoglobulina (anticuerpo) típica comprende un tetrámero. Cada tetrámero está compuesto por dos pares idénticos de cadenas de polipéptido, teniendo cada par una cadena “ligera” (aproximadamente 25 kD) y una “pesada” (aproximadamente 50-70 kD). El extremo N-terminal de cada cadena define una región variable de aproximadamente 100 a 110 o más aminoácidos responsable principalmente del reconocimiento del antígeno. Los términos cadena ligera variable (VL) y cadena pesada variable (VH) se refieren a estas cadenas ligera y pesada respectivamente. Los anticuerpos existen como inmunoglobulinas intactas o como varios fragmentos bien caracterizados producidos por digestión con diversas peptidasas. Por tanto, por ejemplo, la pepsina digiere un anticuerpo por debajo de los enlaces disulfuro en la región de bisagra para producir  $F(ab)'_2$ , un dímero de Fab que por sí mismo es una cadena ligera unida a VH-CH1 mediante un enlace disulfuro. El  $F(ab)'_2$  puede reducirse en condiciones suaves para romper el enlace disulfuro en la región de bisagra convirtiendo de ese modo el dímero  $(Fab)'_2$  en un monómero Fab'. El monómero Fab' es esencialmente un Fab con parte de la región de bisagra (véase, Fundamental Immunology, W.E. Paul, ed., Raven Press, N.Y. (1999), para una descripción más detallada de otros fragmentos de anticuerpo). Aunque se definen diversos fragmentos de anticuerpos en cuanto a la digestión de un anticuerpo intacto, un experto apreciará que tales fragmentos Fab' pueden sintetizarse *de novo* o bien químicamente o bien utilizando metodología de ADN recombinante. Por tanto, el término anticuerpo, tal como se usa en el presente documento, incluye anticuerpos o fragmentos o bien producidos mediante la modificación de anticuerpos completos o bien sintetizados *de novo* usando metodologías de ADN recombinante. Los anticuerpos incluyen, por ejemplo, anticuerpos policlonales, anticuerpos monoclonales, anticuerpos de cadena sencilla o múltiple, incluyendo anticuerpos Fv de cadena sencilla (sFv o scFv) en los que se unen entre sí una cadena pesada variable y una ligera variable (directamente o a través de un ligador peptídico) para formar un polipéptido continuo, y anticuerpos humanizados o quiméricos.

#### Virus influenza

Los polipéptidos y polinucleótidos descritos en el presente documento, por ejemplo, SEC ID N° 1-20, son variantes de secuencias de HA y NA de influenza. En general, los virus influenza están constituidos por un núcleo de ribonucleoproteína interno que contiene un genoma de ARN monocatenario segmentado y una envuelta lipoproteica

externa revestida por una proteína de la matriz. El genoma de virus influenza está compuesto por ocho segmentos de ácido ribonucleico de hebra (-) lineal (ARN), que codifica para las proteínas inmunogénicas hemaglutinina (HA) y neuraminidasa (NA), y seis polipéptidos de núcleo interno: la nucleoproteína de la nucleocápsida (NP); proteínas de la matriz (M); proteínas no estructurales (NS); y 3 proteínas de ARN polimerasa (PA, PB1, PB2). Durante la replicación, el ARN viral genómico se transcribe en ARN mensajero de hebra (+) y ARNc genómico de hebra (-) en el núcleo de la célula huésped. Cada uno de los ocho segmentos genómicos se empaqueta en complejos de ribonucleoproteína que contienen, además del ARN, NP y un complejo de polimerasa (PB1, PB2 y PA).

Influenza se agrupa comúnmente en las categorías influenza A e influenza B. Los virus influenza A e influenza B contienen cada uno ocho segmentos de ARN monocatenario con polaridad negativa. El genoma de influenza A codifica para once polipéptidos. Los segmentos 1-3 codifican para tres polipéptidos, que constituyen una ARN polimerasa dependiente de ARN. El segmento 1 codifica para la proteína PB2 del complejo de polimerasa. Las proteínas de polimerasa restantes PB1 y PA están codificadas por el segmento 2 y el segmento 3, respectivamente. Además, el segmento 1 de algunas cepas de influenza codifica para una proteína pequeña, PB1-F2, producida a partir de un marco de lectura alternativo dentro de la región que codifica para PB1. El segmento 4 codifica para la glicoproteína de superficie hemaglutinina (HA) implicada en la unión y entrada en la célula durante la infección. El segmento 5 codifica para el polipéptido de nucleoproteína de la nucleocápsida (NP), el principal componente estructural asociado con ARN viral. El segmento 6 codifica para una glicoproteína de la envuelta, neuraminidasa (NA). El segmento 7 codifica para dos proteínas de la matriz, denominadas M1 y M2, que se traducen a partir de ARNm cortados y empalmados de manera diferencial. El segmento 8 codifica para NS1 y NS2, dos proteínas no estructurales, que se traducen a partir de variantes de ARNm cortadas y empalmadas de manera alternativa. Los ocho segmentos del genoma de influenza B codifican para 11 proteínas. Los tres genes más grandes codifican para componentes de la ARN polimerasa, PB1, PB2 y PA. El segmento 4 codifica para la proteína HA. El segmento 5 codifica para NP. El segmento 6 codifica para la proteína NA y la proteína NB. Ambas proteínas, NB y NA, se traducen a partir de marcos de lectura solapantes de un ARNm bicistrónico. El segmento 7 de influenza B también codifica para dos proteínas: M1 y BM2. El segmento más pequeño codifica para dos productos: NS1 se traduce a partir del ARN de longitud completa, mientras que NS2 se traduce a partir de una variante de ARNm de corte y empalme.

#### Vacunas contra virus influenza

Las secuencias, composiciones y métodos en el presente documento se refieren principalmente, pero no únicamente, a la producción de virus influenza para vacunas. Históricamente, las vacunas contra el virus influenza se han producido principalmente en huevos de gallina embrionados usando cepas de virus seleccionadas o basadas en predicciones empíricas de cepas relevantes. Más recientemente, se han producido virus reagrupados que incorporan antígenos de hemaglutinina y/o neuraminidasa seleccionados en el contexto de una cepa maestra sensible a la temperatura, atenuada, aprobada. Tras el cultivo del virus a través de múltiples pases en huevos de gallina, se recuperan virus influenza y, opcionalmente, se inactivan, por ejemplo, usando formaldehído y/o propiolactona (o alternativamente se usan en vacunas atenuadas vivas). Por tanto, se apreciará que las secuencias de HA y NA (por ejemplo, SEC ID N° 1-20) son bastante útiles en la construcción de vacunas contra influenza. La presente descripción incluye virus/vacunas que comprenden secuencias de HA y/o NA de cepas de influenza pandémicas (incluyendo cuando las secuencias de HA comprenden sitios de escisión polibásicos modificados tales como las modificaciones descritas en el presente documento); e incluyendo cuando los virus/vacunas comprenden una estructura principal de *ca* tal como A/AA/6/60 o la estructura principal de PR8.

Los intentos de producir vacunas reagrupadas y recombinantes en cultivo celular se han visto dificultados por la incapacidad de algunas de las cepas aprobadas para la producción de vacunas para crecer eficazmente en condiciones de cultivo celular convencionales. Sin embargo, el trabajo previo por los inventores y sus colaboradores proporcionó un sistema de vector y métodos para producir virus reagrupados y recombinantes en cultivo, haciendo posible por tanto producir rápidamente vacunas correspondientes a una o muchas cepas antigénicas seleccionadas de virus, por ejemplo, cepas o bien A o bien B, diversos subtipos o subcepas, etc., por ejemplo, que comprenden las secuencias de HA y/o NA en el presente documento. Véase el sistema de múltiples plásmidos para la producción de virus influenza, documento US 20044029251. Normalmente, los cultivos se mantienen en un sistema, tal como un incubador de cultivo celular, bajo CO<sub>2</sub> y humedad controlados, a temperatura constante usando un regulador de temperatura, tal como un termostato para garantizar que la temperatura no supera 35°C. Pueden obtenerse fácilmente virus influenza reagrupados introduciendo un subconjunto de vectores que corresponden a segmentos genómicos de un virus influenza maestro, en combinación con segmentos complementarios derivados de cepas de interés (por ejemplo, variantes antigénicas de HA y/o NA en el presente documento). Normalmente, las cepas maestras se seleccionan basándose en propiedades deseables relevantes para la administración de vacunas. Por ejemplo, para la producción de vacunas, por ejemplo, para la producción de una vacuna atenuada viva, la cepa de virus donadora maestra puede seleccionarse para un fenotipo atenuado, adaptación al frío y/o sensibilidad a la temperatura. Tal como se explica en otra parte en el presente documento, y, por ejemplo, en el documento US 20044029251, la invención utiliza la cepa de influenza A/Ann Arbor (AA)/6/60 como "estructura principal" sobre la que añadir genes de HA y/o NA (por ejemplo, tal como las secuencias enumeradas en el presente documento, etc.) para crear virus reagrupados deseados tal como se menciona en las reivindicaciones. Por tanto, por ejemplo, en un reagrupamiento 6:2, 2 genes (es decir, NA y HA) serían de la(s) cepa(s) de influenza frente a la(s) que se desea una



reacción inmunogénica, mientras que los otros 6 genes serían de la cepa Ann Arbor, u otra cepa de estructura principal, etc. El virus Ann Arbor es útil por sus atributos adaptados al frío, atenuados y sensibles a la temperatura. Por supuesto, se apreciará que las secuencias de HA y NA en el presente documento pueden reagruparse con varios otros genes de virus o tipos de virus (por ejemplo, varias “estructuras principales” diferentes tales como PR8, etc., que contienen los otros genes de influenza presentes en un reagrupamiento, concretamente, los genes que no son HA ni NA).

Diversas realizaciones en el presente documento pueden comprender vacunas atenuadas vivas tal como se menciona en las reivindicaciones, que tienen las secuencias de HA y NA que comprenden respectivamente SEC ID N° 11 y SEC ID N°. 12 en el presente documento, para gripe pandémica. Un problema que surge del crecimiento de cepas de virus de vacuna (por ejemplo, reagrupados) en huevos es que las cepas aviares (que pueden estar implicadas en pandemias) pueden destruir los huevos en los que van a producirse las vacunas y, por tanto, son difíciles de manipular, producir, etc. a través del uso de producción reagrupada tradicional (sin rescate con plásmidos). Tales cepas aviares son de interés puesto que las pruebas indican que pueden dar como resultado gripe en seres humanos y posibles pandemias. Por tanto, el uso de sistemas de rescate con plásmidos para crear/manipular reagrupados de influenza con cepas pandémicas tales como diversas secuencias aviares (por ejemplo, las secuencias de HA y NA en el presente documento) son bastante deseables. Se apreciará, sin embargo, que las secuencias actuales también pueden usarse con sistemas no plasmídicos o tradicionales.

Pueden infectarse aves acuáticas (entre otras) por virus influenza A de los subtipos hemaglutinina 15 (HA) y neuraminidasa 9 (NA). Tales aves pueden servir como reservorio a partir del que pueden introducirse subtipos de influenza novedosos en poblaciones humanas y provocar pandemias. La observación de que virus influenza A H7N7 aviares infectaron a seres humanos en los Países Bajos en 2003 y que virus H5N1 y H9N2 aviares infectaron a seres humanos en Hong Kong y China anteriormente, suscita la inquietud de que estos (y otros) subtipos tengan el potencial de provocar pandemias. Por tanto, se necesitan vacunas para prevenir infecciones humanas con virus influenza A aviares. Recientemente, se autorizaron en los Estados Unidos vacunas de virus influenza A atenuado, vivo contra virus influenza humanos. Véase anteriormente. Tales vacunas son virus H1N1 y H3N2 reagrupados en los que los genes de proteínas internos de virus adaptado al frío (*ca*) de A/Ann Arbor (AA)/6/60 (H2N2) confieren los fenotipos adaptado al frío, de atenuación y sensible a la temperatura de virus AA *ca* en los virus reagrupados (es decir, los que tienen los genes de hemaglutinina y neuraminidasa de la cepa distinta de Ann Arbor). Se han aplicado técnicas de genética inversa basada en plásmidos y reagrupamiento genético clásico para generar virus reagrupados que contienen los genes de hemaglutinina y neuraminidasa de virus influenza A aviares (subtipos H4-H14) y seis segmentos génicos internos del virus AA *ca*. Tales virus reagrupados tal como se menciona en las reivindicaciones son características de la invención. Véanse las secuencias de genes de HA y NA a continuación. Estos virus portan propiedades biológicas que son deseables en vacunas candidatas porque los fenotipos asociados con los virus AA *ca* están presentes en los virus reagrupados. La generación y evaluación de estos virus reagrupados como virus simiente para vacunas son etapas importantes en la preparación pandémica. Se contempla que ensayos clínicos puedan establecer la seguridad, infectividad e inmunogenicidad de tales vacunas pandémicas atenuadas vivas. También se incluyen métodos de construcción y uso de tales virus y vacunas. “Cepas de virus pandémicas” tal como se usa en el presente documento se define como un subtipo de virus influenza cepa A que no está circulando en la población humana, que se declara que es una cepa pandémica por los Centros para el Control de Enfermedades o la Organización Mundial de la Salud o se reconocen generalmente como tales dentro de la comunidad científica.

Tal como se describe en el presente documento, las secuencias antigénicas (por ejemplo, las secuencias de HA) así como virus y vacunas de tales virus comprenden sitios de escisión polibásicos modificados. Algunas cepas de influenza pandémicas aviares altamente patogénicas comprenden múltiples sitios de escisión de aminoácidos básicos dentro de secuencias de hemaglutinina. Véase, por ejemplo, Li *et al.*, J. of Infectious Diseases, 179:1132-8, 1999. Tales sitios de escisión, en realizaciones típicas en el presente documento, por ejemplo, están modificados o alterados en sus secuencias en comparación con las secuencias de tipo natural de las que se derivan las secuencias actuales (por ejemplo, para inutilizar la escisión o reducir la escisión en las mismas, etc.). Tales modificaciones/alteraciones pueden ser diferentes en diferentes cepas debido a las diversas secuencias de los sitios de escisión en las secuencias de tipo natural. Por ejemplo, se eliminan normalmente en secuencias en el presente documento 4 residuos polibásicos (RRKK) en 326-329 de H5 madura (en comparación con wt). Véase la lista de secuencias y la figura 1. Los sitios de escisión polibásicos pueden modificarse de varios modos. Por ejemplo, el sitio de escisión polibásico puede eliminar un aminoácido de una vez (por ejemplo un R eliminado, dos R eliminados, RRK eliminados o RRKK eliminados). Adicionalmente, el residuo de aminoácido directamente en el sentido de 5' del sitio de escisión también puede eliminarse o alterarse (por ejemplo, de un R a un T, etc.); además, también pueden modificarse los nucleótidos que codifican para el residuo de aminoácido directamente tras el sitio de escisión. Véase, por ejemplo, la figura 1 para una ilustración de la modificación del sitio de escisión. Además, secuencias de polipéptido de hemaglutinina de virus influenza comprenden secuencias de péptido señal amino terminal, por tanto, las secuencias de polipéptido de hemaglutinina incluyen tanto la forma madura (péptido señal amino terminal escindido) de los polipéptidos de hemaglutinina como la forma pre-escindida de hemaglutinina. Los sitios de escisión de cualquier secuencia de polipéptido de hemaglutinina de cualquier cepa de influenza pueden medirse o predecirse rutinariamente usando cualquiera de varios métodos en la técnica.

Los términos “sensible a la temperatura”, “adaptado al frío” y “atenuado” tal como se aplican a virus (normalmente usados como vacunas o para la producción de vacunas) que abarcan opcionalmente las presentes secuencias, se conocen bien en la técnica. Por ejemplo, el término “sensible a la temperatura” (*ts*) indica, por ejemplo, que el virus presenta una reducción de 100 veces o más en el título a 39°C en relación con 33°C para cepas de influenza A, o que el virus presenta una reducción de 100 veces o más en el título a 37°C en relación con 33°C para cepas de influenza B. El término “adaptado al frío” (*ca*) indica que el virus presenta crecimiento a 25°C dentro de 100 veces de su crecimiento a 33°C, mientras que el término “atenuado” (*att*) indica que el virus se replica en las vías respiratorias superiores de hurones pero no es detectable en sus tejidos pulmonares, y no provoca enfermedad de tipo gripe en el animal. Se entenderá que virus con fenotipos intermedios, es decir, virus que presentan reducciones en el título inferiores a 100 veces a 39°C (para virus de la cepa A) o 37°C (para virus de la cepa B), o que presentan crecimiento a 25°C que es mayor de 100 veces su crecimiento a 33°C (por ejemplo, dentro de 200 veces, 500 veces,

1000 veces, 10.000 veces menos), y/o presentan crecimiento reducido en los pulmones en relación con el crecimiento en las vías respiratorias superiores de hurones (es decir, parcialmente atenuados) y/o enfermedad de tipo gripe en el animal, también son virus útiles y pueden usarse conjuntamente con las secuencias de HA y NA en el presente documento.

De nuevo, las secuencias de HA y NA descritas en el presente documento se utilizan en tales vacunas reagrupadas con plásmidos (y/o en otros virus y vacunas *ts*, *cs*, *ca* y/o *att*).

#### Flumist™

Tal como se mencionó anteriormente, existen numerosos ejemplos y tipos de vacunas contra influenza. Una vacuna contra influenza a modo de ejemplo es FluMist™ que es una vacuna viva, atenuada que protege a niños y adultos de la enfermedad de la gripe (Belshe *et al.* (1998) The efficacy of live attenuated, cold-adapted, trivalent, intranasal influenza virus vaccine in children N Engl J Med 338:1405-12; Nichol *et al.* (1999) Effectiveness of live, attenuated intranasal influenza virus vaccine in healthy, working adults: a randomized controlled trial JAMA 282:137-44).

Las cepas de la vacuna FluMist™ contienen, por ejemplo, segmentos de genes de HA y NA derivados de las cepas (por ejemplo, cepas de tipo natural) a las que se dirige la vacuna junto con seis segmentos de genes, PB1, PB2, PA, NP, M y NS, de un virus donador maestro común (MDV). El MDV para las cepas de influenza A de FluMist (MDV-A) se creó mediante pase en serie de la cepa A/Ann Arbor/6/60 (A/AA/6/60) de tipo natural en cultivo tisular de riñón de pollo primario a temperaturas sucesivamente inferiores (Maassab (1967) Adaptation and growth characteristics of influenza virus at 25 degrees C Nature 213:612-4). MDV-A se replica eficazmente a 25°C (*ca*, adaptado al frío), pero su crecimiento se restringe a 38 y 39°C (*ts*, sensible a la temperatura). Adicionalmente, este virus no se replica en los pulmones de hurones infectados (*att*, atenuación). Se cree que el fenotipo *ts* contribuye a la atenuación de la vacuna en seres humanos restringiendo su replicación en todas excepto las regiones más frías del aparato respiratorio. La estabilidad de esta propiedad se ha demostrado en modelos animales y estudios clínicos. En contraposición al fenotipo *ts* de cepas de influenza creadas mediante mutagénesis química, la propiedad *ts* de MDV-A no se revierte tras el pase a través de hámsteres infectados o en aislados extraídos de niños (para una revisión reciente, véase Murphy & Coelingh (2002) Principles underlying the development and use of live attenuated cold-adapted influenza A y B virus vaccines Viral Immunol 15:295-323).

Estudios clínicos en más de 20.000 adultos y niños que implicaban 12 cepas con reagrupamiento 6:2 separadas han mostrado que estas vacunas están atenuadas, son seguras y eficaces (Belshe *et al.* (1998) The efficacy of live attenuated, cold-adapted, trivalent, intranasal influenza virus vaccine in children N Engl J Med 338:1405-12; Boyce *et al.* (2000) Safety and immunogenicity of adjuvanted and unadjuvanted subunit influenza vaccines administered intranasally to healthy adults Vaccine 19:217-26; Edwards *et al.* (1994) A randomized controlled trial of cold adapted and inactivated vaccines for the prevention of influenza A disease J Infect Dis 169:68-76; Nichol *et al.* (1999) Effectiveness of live, attenuated intranasal influenza virus vaccine in healthy, working adults: a randomized controlled trial JAMA 282:137-44). Los reagrupados que portan los seis genes internos de MDV-A y los dos segmentos de genes de HA y NA de un virus de tipo natural (es decir, un reagrupamiento 6:2) mantienen de manera constante los fenotipos *ca*, *ts* y *att* (Maassab *et al.* (1982) Evaluation of a cold-recombinant influenza virus vaccine in ferrets J. Infect. Dis. 146:780-900).

La producción de tales virus reagrupados usando cepas B de influenza es más difícil, sin embargo, el trabajo reciente (véase, por ejemplo, Sistema de múltiples plásmidos para la producción de virus Influenza, documento US 20044029251) ha mostrado un sistema de ocho plásmidos para la generación de virus influenza B completamente a partir de ADNc clonado. También se mostraron métodos para la producción de virus influenza A y B vivos atenuados adecuados para formulaciones de vacuna, tales como formulaciones de vacuna de virus vivos útiles para administración intranasal.

El sistema y los métodos descritos anteriormente son útiles para la producción rápida en cultivo celular de virus influenza A y B recombinantes y reagrupados, incluyendo virus adecuados para su uso como vacunas, incluyendo vacunas atenuadas vivas, tales como vacunas adecuadas para administración intranasal. Las secuencias (por ejemplo, SEC ID N° 1-10 y los aminoácidos correspondientes en SEC ID N° 11-20), los métodos, etc. descritos en el

presente documento opcionalmente se usan conjuntamente con, o en combinación con, tal trabajo previo que implica, por ejemplo, virus influenza reagrupados para la producción de vacunas para producir virus para vacunas.

#### Métodos y composiciones para la administración profiláctica de vacunas

Los virus recombinantes y reagrupados descritos en el presente documento (por ejemplo, los que comprenden los polinucleótidos de SEC ID N° 1-10 o los polipéptidos de SEC ID N° 11-20, o fragmentos de los mismos), en particular los citados en las reivindicaciones, pueden administrarse profilácticamente en una cantidad inmunológicamente eficaz y en un vehículo o excipiente apropiado para estimular una respuesta inmunitaria específica para una o más cepa(s) de virus influenza tal como se determina mediante la secuencia de HA y/o NA. Normalmente, el vehículo o excipiente es un vehículo o excipiente farmacéuticamente aceptable, tal como agua estéril, solución salina acuosa, soluciones salinas tamponadas acuosas, disoluciones de dextrosa acuosas, disoluciones de glicerol acuosas, etanol o combinaciones de los mismos. La preparación de tales disoluciones que garantizan la esterilidad, el pH, la isotonicidad y la estabilidad se efectúa según protocolos establecidos en la técnica. Generalmente, se selecciona un vehículo o excipiente para minimizar efectos alérgicos y otros efectos no deseados, y para adecuarse a la vía de administración particular, por ejemplo, subcutánea, intramuscular, intranasal, etc.

Un aspecto relacionado de la invención proporciona el virus influenza reagrupado tal como se menciona en las reivindicaciones para su uso en métodos para estimular el sistema inmunitario de un individuo para producir una respuesta inmunitaria protectora frente al virus influenza. En los métodos, se administra una cantidad inmunológicamente eficaz de un virus influenza recombinante (por ejemplo, que comprende una molécula de HA y/o NA tal como se describe en el presente documento) al individuo en un vehículo fisiológicamente aceptable.

Generalmente, los virus influenza de la invención se administran en una cantidad suficiente para estimular una respuesta inmunitaria específica para una o más cepa(s) de virus influenza (es decir, contra las cepas de HA y/o NA descritas en el presente documento). Preferiblemente, la administración de los virus influenza provoca una respuesta inmunitaria protectora frente a tales cepas. Los expertos en la técnica conocen dosificaciones y métodos para provocar una respuesta inmunitaria protectora contra una o más cepa(s) de influenza. Véanse, por ejemplo, la patente de EE.UU. n° 5.922.326; Wright *et al.*, Infect. Immun. 37:397-400 (1982); Kim *et al.*, Pediatrics 52:56-63 (1973); y Wright *et al.*, J. Pediatr. 88:931-936 (1976). Por ejemplo, se proporcionan virus influenza en el intervalo de aproximadamente 1-1000 HID50 (dosis infecciosa humana), es decir, aproximadamente 105 - 108 ufp (unidades formadoras de placa) por dosis administrada. Normalmente, la dosis se ajustará dentro de este intervalo basándose, por ejemplo, en la edad, el estado físico, el peso corporal, el sexo, la dieta, el momento de administración y otros factores clínicos. La formulación de vacuna profiláctica se administra sistémicamente, por ejemplo, mediante inyección subcutánea o intramuscular usando una aguja y jeringuilla, o un dispositivo de inyección sin aguja. Alternativamente, la formulación de vacuna se administra por vía intranasal, mediante o bien gotas, aerosol de partículas grandes (mayores de aproximadamente 10 micrómetros) o bien pulverización en el aparato respiratorio superior. Aunque cualquiera de las vías anteriores de administración da como resultado una respuesta inmunitaria sistémica protectora, la administración intranasal confiere el beneficio añadido de provocar inmunidad mucosa en el sitio de entrada del virus influenza. Para la administración intranasal, se prefieren a menudo vacunas de virus vivos atenuados, por ejemplo, un virus influenza recombinante o reagrupado atenuado, adaptado al frío y/o sensible a la temperatura. Véase anteriormente. Aunque se prefiere la estimulación de una respuesta inmunitaria protectora con una única dosis, pueden administrarse dosificaciones adicionales, mediante la misma vía o diferente, para lograr el efecto profiláctico deseado.

Normalmente, el virus influenza recombinante atenuado de la presente invención tal como se usa en una vacuna está suficientemente atenuado de manera que no se producirán síntomas de infección, o al menos síntomas de infección grave, en la mayoría de los individuos inmunizados (o infectados de otra forma) con el virus influenza atenuado. En algunos casos, el virus influenza atenuado todavía puede producir síntomas de enfermedad leve (por ejemplo, enfermedad respiratoria superior leve) y/o diseminación a individuos no vacunados. Sin embargo, su virulencia está lo suficientemente suprimida de manera que no se producen infecciones del aparato respiratorio inferior graves en el huésped vacunado o accidental.

Alternativamente, puede estimularse una respuesta inmunitaria seleccionando como diana *ex vivo* o *in vivo* células dendríticas con virus influenza que comprenden las secuencias en el presente documento. Por ejemplo, se exponen células dendríticas en proliferación a virus en una cantidad suficiente y durante un periodo de tiempo suficiente para permitir la captura de los antígenos de influenza por las células dendríticas. Las células se transfieren entonces a un sujeto que va a vacunarse mediante métodos de trasplante intravenoso convencional.

Aunque se prefiere la estimulación de una respuesta inmunitaria protectora con una única dosis, pueden administrarse dosificaciones adicionales, mediante la misma o diferente vía, para lograr el efecto profiláctico deseado. En neonatos y lactantes, por ejemplo, pueden requerirse múltiples administraciones para provocar niveles suficientes de inmunidad. La administración puede continuar a intervalos durante toda la infancia, según sea necesario para mantener niveles suficientes de protección frente a la infección por influenza de tipo natural. De manera similar, adultos que son particularmente susceptibles a infección por influenza repetida o grave, tales como, por ejemplo, trabajadores de atención sanitaria, trabajadores de guarderías, miembros de la familia de niños

jóvenes, los ancianos e individuos con función cardiopulmonar comprometida pueden requerir múltiples inmunizaciones para establecer y/o mantener respuestas inmunitarias protectoras. Los niveles de inmunidad inducida pueden monitorizarse, por ejemplo, midiendo las cantidades de anticuerpos séricos y secretores neutralizantes, y ajustarse las dosificaciones o repetirse las vacunaciones según sea necesario para provocar y

5 mantener niveles deseados de protección.

Opcionalmente, la formulación para administración profiláctica de los virus influenza también contiene uno o más adyuvantes para potenciar la respuesta inmunitaria frente a los antígenos de influenza. Los adyuvantes adecuados incluyen: adyuvante completo de Freund, adyuvante incompleto de Freund, saponina, geles minerales tales como

10 hidróxido de aluminio, sustancias tensioactivas tales como lisolectina, polioles plurónicos, polianiones, péptidos, emulsiones en aceite o hidrocarburos, bacilo Calmette-Guerin (BCG), *Corynebacterium parvum* y el adyuvante sintético QS-21.

15 Si se desea, puede realizarse la administración de vacuna profiláctica de virus influenza conjuntamente con la administración de una o más moléculas inmunoestimuladoras. Las moléculas inmunoestimuladoras incluyen diversas citocinas, linfocinas y quimiocinas con actividades inmunoestimuladoras, inmunopotenciadoras y proinflamatorias, tales como interleucinas (por ejemplo, IL-1, IL-2, IL-3, IL-4, IL-12, IL-13); factores de crecimiento (por ejemplo, factor estimulante (CSF) de colonias de granulocitos-macrófagos (GM)); y otras moléculas

20 inmunoestimuladoras, tales como tal como factor inflamatorio de macrófagos, ligando Flt3, B7.1; B7.2, etc. Las moléculas inmunoestimuladoras pueden administrarse en la misma formulación que los virus influenza, o pueden administrarse por separado.

Aunque la vacunación de un individuo con un virus de influenza atenuado de una cepa particular de un subgrupo particular puede inducir protección cruzada contra virus influenza de diferentes cepas y/o subgrupos, la protección

25 cruzada puede potenciarse, si se desea, vacunando al individuo con virus influenza atenuados de al menos dos cepas, por ejemplo, cada una de las cuales representa un subgrupo diferente. Adicionalmente, las combinaciones de vacunas pueden incluir opcionalmente mezclas de vacunas pandémicas (por ejemplo, aquellas contra cepas de influenza pandémicas tales como diversas cepas aviáres, véanse, por ejemplo, las secuencias en el presente documento, u otras cepas pandémicas) y cepas no pandémicas. Las mezclas de vacunas (o vacunaciones múltiples)

30 pueden comprender componentes de cepas humanas y/o cepas de influenza no humanas (por ejemplo, aviáres y humanas, etc.). De manera similar, las vacunas de virus influenza atenuados de la presente invención pueden combinarse opcionalmente con vacunas que inducen respuestas inmunitarias protectoras contra otros agentes infecciosos.

### 35 Polinucleótidos

Se sabe bien en la técnica que los segmentos de polinucleótidos de HA y NA de virus influenza comprenden tanto una región codificante (que codifica para el ORF) como regiones no codificantes (NCR), tanto en 5' como en 3' de la

40 secuencia que codifica para HA y NA. Se muestran ejemplos de estas NCR en SEC ID N° 1-9 (fuera de los ORF). También se sabe que pueden prepararse cebadores para estas NCR para facilitar la amplificación de los segmentos de HA y NA completos de virus influenza. (Véase, por ejemplo, Hoffmann *et al.* Arch Virol. Diciembre de 2001;

146(12): 2275-89). Además, se sabe que las NCR de la HA y NA de influenza pueden aumentar la eficacia de lograr reagrupamientos. Por tanto, las secuencias de polinucleótido de estas NCR se describen en el presente documento.

45 Cuando se amplifican los segmentos de HA y NA de cualquiera cepa pandémica, podrían prepararse y usarse cebadores de polinucleótido para unirse a regiones conservadas (por ejemplo, entre cepas relacionadas) de las NCR de HA y NA para la amplificación (por ejemplo, mediante RT-PCR).

Los polinucleótidos de HA y NA del virus de la invención, SEC ID N° 1 o por ejemplo SEC ID N° 2, se usan

50 opcionalmente en varias capacidades diferentes alternativamente a, o además de, las vacunas descritas anteriormente. Se describen en el presente documento otros usos a modo de ejemplo en el presente documento para fines ilustrativos y no como limitaciones en la gama actual de usos. También se describen en el presente documento diferentes métodos de construcción, purificación y caracterización de las secuencias de nucleótidos de la descripción. Se describe en el presente documento que ácidos nucleicos incluyendo una o más secuencia(s) de

55 polinucleótido(s) descritas anteriormente se usan favorablemente como sondas para la detección de ácidos nucleicos correspondientes o relacionados en una variedad de contextos, tal como en experimentos de hibridación de ácidos nucleicos, por ejemplo, para encontrar y/o caracterizar variantes de influenza homólogas (por ejemplo, homólogos para las secuencias en el presente documento, etc.) que infectan a otras especies o en brotes de gripe diferentes, etc. Las sondas pueden ser moléculas de o bien ADN o bien ARN, tales como fragmentos de restricción

60 de ADN genómico o clonado, ADNc, productos de amplificación por PCR, transcritos y oligonucleótidos, y pueden variar en longitud desde oligonucleótidos tan cortos como aproximadamente 10 nucleótidos de longitud hasta secuencias de longitud completa o ADNc en exceso de 1 kb o más. Por ejemplo, una sonda incluye una secuencia o subsecuencia de polinucleótido seleccionada, por ejemplo, de entre SEC ID N° 1 o por ejemplo SEC ID N° 2, o secuencias complementarias a las mismas. Alternativamente, se usan como sondas secuencias de polinucleótido

65 que son variantes de una de las secuencias designadas anteriormente. Lo más normalmente, tales variantes incluyen una o unas cuantas variaciones de nucleótidos conservadoras. Por ejemplo, pueden seleccionarse pares (o

conjuntos) de oligonucleótidos, en los que las dos (o más) secuencias de polinucleótido son variaciones conservadoras entre sí, en las que una secuencia de polinucleótido corresponde idénticamente a una primera variante o/y la(s) otra(s) corresponden idénticamente a variantes adicionales. Tales pares de sondas oligonucleotídicas son particularmente útiles, por ejemplo, para experimentos de hibridación específicos para detectar nucleótidos polimórficos o, por ejemplo, para detectar variantes de HA y NA de influenza homólogas, por ejemplo, homólogas a las secuencias de HA y NA actuales, que infectan a otras especies o están presentes en brotes de gripe diferentes (por ejemplo, o bien temporalmente y/o bien geográficamente diferentes). En otras aplicaciones, se seleccionan sondas que son más divergentes, es decir, se seleccionan sondas que son al menos aproximadamente el 91% (o aproximadamente el 92%, aproximadamente el 93%, aproximadamente el 94%, aproximadamente el 95%, aproximadamente el 96%, aproximadamente el 97%, aproximadamente el 98%, aproximadamente el 98,5%, aproximadamente el 98,7%, aproximadamente el 99%, aproximadamente el 99,1%, aproximadamente el 99,2%, aproximadamente el 99,3%, aproximadamente el 99,4%, aproximadamente el 99,5%, o aproximadamente el 99,6% o más aproximadamente el 99,7%, aproximadamente el 99,8%, aproximadamente el 99,9% o más) idénticas.

Las sondas descritas en el presente documento por ejemplo, tal como se muestra modo de ejemplo mediante secuencias derivadas de las secuencias en el presente documento, también pueden usarse para identificar secuencias de polinucleótido útiles adicionales según procedimientos rutinarios en la técnica. Se utilizan una o más sondas, tal como se describió anteriormente, para examinar bibliotecas de productos de expresión o segmentos cromosómicos (por ejemplo, bibliotecas de expresión o bibliotecas genómicas) para identificar clones que incluyen secuencias idénticas a, o con similitud de secuencia significativa con, por ejemplo, una o más sonda(s) de las secuencias en el presente documento, es decir, variantes, homólogos, etc. Se entenderá que además de métodos físicos tales como examen de bibliotecas, también pueden usarse enfoques bioinformáticos asistidos por ordenador, por ejemplo, BLAST y otros algoritmos de búsqueda de homología de secuencia, y similares, para identificar secuencias de polinucleótido relacionadas.

Se producen opcionalmente sondas oligonucleotídicas mediante una variedad de métodos bien conocidos por los expertos en la técnica. Lo más normalmente, se producen mediante métodos sintéticos bien conocidos, tales como el método de triéster de fosforamídita en fase sólida descrito por Beaucage y Caruthers (1981) *Tetrahedron Letts* 22(20): 1859-1862, por ejemplo, usando un sintetizador automatizado, o tal como se describe en Needham-Van Devanter *et al.* (1984) *Nucl Acids Res*, 12: 6159-6168. También pueden prepararse oligonucleótidos a medida y pedirse de una variedad de fuentes comerciales conocidas por los expertos. La purificación de oligonucleótidos, cuando sea necesario, se realiza normalmente mediante o bien electroforesis en gel de acrilamida nativa o bien mediante HPLS de intercambio aniónico tal como se describe en Pearson y Regnier (1983) *J Chrom* 255: 137-149. La secuencia de los oligonucleótidos sintéticos puede verificarse usando el método de degradación química de Maxam y Gilbert (1980) en Grossman and Moldave (eds.) *Academic Press*, Nueva York, *Methods in Enzymology* 65:499-560. También pueden pedirse fácilmente oligonucleótidos a medida de una variedad de fuentes comerciales conocidas por los expertos.

En otras circunstancias, por ejemplo, que se refieren a atributos de células u organismos que expresan los polinucleótidos y polipéptidos descritos en el presente documento (por ejemplo, los que albergan virus que comprenden las secuencias descritas en el presente documento), se utilizan favorablemente sondas que son polipéptidos, péptidos o anticuerpos. Por ejemplo, se usan favorablemente polipéptidos aislados o recombinantes, fragmentos de polipéptido y péptidos derivados de cualquiera de las secuencias de aminoácidos descritas en el presente documento (por ejemplo, SEC ID N° 11-12) y/o codificados por secuencias de polinucleótido descritas en el presente documento, por ejemplo, seleccionadas de SEC ID N° 1 a SEC ID N° 10, para identificar y aislar anticuerpos, por ejemplo, a partir de bibliotecas de presentación en fago, bibliotecas combinatorias, sueros policlonales y similares. Los fragmentos de polipéptido incluyen un péptido o polipéptido que comprende una secuencia de aminoácidos de al menos 5 residuos de aminoácido contiguos, o al menos 10 residuos de aminoácido contiguos, o al menos 15 residuos de aminoácido contiguos, o al menos 20 residuos de aminoácido contiguos, o al menos 25 residuos de aminoácido contiguos, o al menos 40 residuos de aminoácido contiguos, o al menos 50 residuos de aminoácido contiguos, o al menos 60 residuos de aminoácido contiguos, o al menos 70 residuos de aminoácido contiguos, o al menos 80 residuos de aminoácido contiguos, o al menos 90 residuos de aminoácido contiguos, o al menos 100 residuos de aminoácido contiguos, o al menos 125 residuos de aminoácido contiguos, o al menos 150 residuos de aminoácido contiguos, o al menos 175 residuos de aminoácido contiguos, o al menos 200 residuos de aminoácido contiguos, o al menos 250 residuos de aminoácido contiguos, o al menos 350 residuos de aminoácido contiguos, o al menos 400 contiguos, o al menos 450 contiguos, o al menos 500 contiguos, o al menos 550 contiguos de la secuencia de aminoácidos de un polipéptido de HA o NA descrito en el presente documento (por ejemplo, SEC ID N° 11-20). También se describen en el presente documento polinucleótidos que codifican para dichos fragmentos de polipéptido y anticuerpos que se unen específicamente a dichos polipéptidos.

Anticuerpos específicos para cualquier secuencia o subsecuencia de polipéptido, por ejemplo, de SEC ID N° 11 a SEC ID N° 20, y/o codificados por secuencias de polinucleótido descritas en el presente documento, por ejemplo, seleccionadas de SEC ID N° 1 a SEC ID N° 10, son así mismo valiosos como sondas para evaluar productos de expresión, por ejemplo, a partir de células o tejidos. Además, los anticuerpos son particularmente adecuados para evaluar la expresión de proteínas que comprenden subsecuencias de aminoácidos, por ejemplo, de las que se

facilitan en el presente documento, o codificadas por secuencias de polinucleótido descritas en el presente documento, por ejemplo, seleccionadas de las mostradas en el presente documento, *in situ*, en un alineamiento de tejido, en una célula, tejido u organismo, por ejemplo, un organismo infectado por un virus influenza no identificado o similar. Los anticuerpos pueden marcarse directamente con un reactivo detectable, o detectarse indirectamente marcando un anticuerpo secundario específico para la región constante de cadena pesada (es decir, isotipo) del anticuerpo específico). Se proporcionan a continuación detalles adicionales referentes a la producción de anticuerpos específicos.

#### Ensayos de diagnóstico

Las secuencias de ácido nucleico descritas en el presente documento pueden usarse en ensayos de diagnóstico para detectar influenza (y/o hemaglutinina y/o neuraminidasa) en una muestra, para detectar secuencias similares a hemaglutinina y/o similares a neuraminidasa, y para detectar diferencias de cepas en aislados clínicos de influenza usando fragmentos de polinucleótido o bien químicamente sintetizados o bien recombinantes, por ejemplo, seleccionados de las secuencias en el presente documento. Por ejemplo, pueden usarse fragmentos de las secuencias de hemaglutinina y/o neuraminidasa que comprenden al menos entre 10 y 20 nucleótidos como cebadores para amplificar ácidos nucleicos usando métodos de reacción en cadena de la polimerasa (PCR) bien conocidos en la técnica (por ejemplo, PCR con transcripción inversa) y como sondas en ensayos de hibridación de ácidos nucleicos para detectar material genético diana tal como ARN de influenza en muestras clínicas.

Las sondas descritas en el presente documento, por ejemplo, tal como se muestra a modo de ejemplo mediante subsecuencias únicas seleccionadas de las facilitadas en el presente documento, también pueden usarse para identificar secuencias de polinucleótido útiles adicionales (tal como para caracterizar cepas de influenza adicionales) según procedimientos de rutina en la técnica. Una o más sondas, tal como se describió anteriormente, se utilizan para examinar bibliotecas de productos de expresión o ácidos nucleicos virales clonados (es decir, bibliotecas de expresión o bibliotecas genómicas) para identificar clones que incluyen secuencias idénticas a, o con identidad de secuencia significativa con las secuencias en el presente documento. A su vez, cada una de estas secuencias identificadas puede usarse para preparar sondas, incluyendo pares o conjuntos de sondas variantes tal como se describió anteriormente. Se entenderá que además de tales métodos físicos tales como examen de bibliotecas, también pueden usarse enfoques bioinformáticos asistidos por ordenador, por ejemplo, BLAST y otros algoritmos de búsqueda de homología de secuencia, y similares, para identificar secuencias de polinucleótido relacionadas.

Las sondas descritas en el presente documento son particularmente útiles para detectar la presencia y para determinar la identidad de ácidos nucleicos de influenza en células, tejidos u otras muestras biológicas (por ejemplo, un lavado nasal o lavado bronquial). Por ejemplo, las sondas descritas en el presente documento se utilizan favorablemente para determinar si una muestra biológica, tal como un sujeto (por ejemplo, un sujeto humano) o sistema modelo (tal como una muestra celular cultivada) se ha expuesto a, o se ha infectado con influenza, o cepa(s) particular(es) de influenza. La detección de la hibridación de la sonda seleccionada con ácidos nucleicos que se originan en (por ejemplo, aislados de) la muestra biológica o sistema modelo es indicativa de exposición a o infección con el virus (o un virus relacionado) del que se selecciona el polinucleótido sonda.

Se apreciará que el diseño de la sonda se ve influido por la aplicación prevista. Por ejemplo, cuando van a detectarse varias interacciones sonda-diana específicas de alelo en un único ensayo, por ejemplo, en un único chip de ADN, es deseable tener temperaturas de fusión similares para todas las sondas. Por consiguiente, las longitudes de las sondas se ajustan de modo que las temperaturas de fusión para todas las sondas en el alineamiento sean estrechamente similares (se apreciará que pueden necesitarse diferentes longitudes para diferentes sondas para lograr una  $T_m$  particular cuando diferentes sondas tienen diferentes contenidos en GC). Aunque la temperatura de fusión es una consideración primaria en el diseño de la sonda, se usan opcionalmente otros factores para ajustar adicionalmente la construcción de sondas, tales como seleccionar frente a la autocomplementariedad de cebadores y similares.

#### Vectores, promotores y sistemas de expresión

También se describen en el presente documento construcciones recombinantes que incorporan una o más de las secuencias de ácido nucleico descritas en el presente documento. Tales construcciones incluyen opcionalmente un vector, por ejemplo, un plásmido, un cósmido, un fago, un virus, un cromosoma artificial bacteriano (BAC), un cromosoma artificial de levadura (YAC), etc., en el que una o más de las secuencias de polinucleótido de SEC ID N° 1 o por ejemplo, SEC ID N° 2, o una subsecuencia de las mismas, etc., se ha insertado, en una orientación directa o inversa. Por ejemplo, el ácido nucleico insertado puede incluir una secuencia cromosómica viral o ADNc que incluye toda o parte de al menos una de las secuencias de polinucleótido descritas en el presente documento. En un aspecto, el constructo comprende además secuencias reguladoras, incluyendo, por ejemplo, un promotor, operativamente unido a la secuencia. Los expertos en la técnica conocen grandes números de promotores y vectores adecuados, y están disponibles comercialmente.

Los polinucleótidos descritos en el presente documento pueden incluirse en uno cualquiera de una variedad de vectores adecuados para generar ARN sentido o antisentido, y opcionalmente, productos de expresión de

polipéptido (o péptido) (por ejemplo, una molécula de hemaglutinina y/o neuraminidasa tal como se describe en el presente documento). Tales vectores incluyen secuencias de ADN cromosómico, no cromosómico y sintético, por ejemplo, derivados de SV40; plásmidos bacterianos; ADN de fago; baculovirus; plásmidos de levadura; vectores derivados de combinaciones de plásmidos y ADN de fago, ADN viral tal como vaccinia, adenovirus, virus de viruela aviar, pseudorrabia, adenovirus, virus adenoasociados, retrovirus y muchos otros (por ejemplo, pCDL). Puede usarse cualquier vector que pueda introducir material genético en una célula y, si se desea replicación, que pueda replicarse en el huésped relevante.

En un vector de expresión, la secuencia de polinucleótido de HA y/o NA de interés está físicamente dispuesta en proximidad y orientación con una secuencia de control de la transcripción apropiada (por ejemplo, promotor, y opcionalmente uno o más potenciadores) para dirigir la síntesis de ARNm. Es decir, la secuencia de polinucleótido de interés está operativamente unida a una secuencia de control de la transcripción apropiada. Los ejemplos de tales promotores incluyen: promotor de SV40 o LTR, promotor lac o trp de *E. coli*, promotor PL de fago lambda y otros promotores que se sabe que controlan la expresión de genes en células procariotas o eucariotas y sus virus.

Una variedad de promotores son adecuados para su uso en vectores de expresión para regular la transcripción de segmentos de genoma de virus influenza. En determinadas realizaciones, se utiliza el promotor de la ARN polimerasa II (Pol II) dependiente de ADN de citomegalovirus (CMV). Si se desea, por ejemplo, para regular la expresión condicional, pueden sustituirse otros promotores que inducen transcripción de ARN en las condiciones especificadas, o en las células o tejidos especificados. Están disponibles numerosos promotores virales y de mamíferos, por ejemplo, humanos, o pueden aislarse según la aplicación específica contemplada. Por ejemplo, los promotores alternativos obtenidos de los genomas de virus animales y humanos incluyen promotores tales como el adenovirus (tal como adenovirus 2), virus del papiloma, virus de la hepatitis B, virus del poliovirus y virus del simio 40 (SV40), y diversos promotores retrovirales. Los promotores de mamíferos incluyen, entre muchos otros, el promotor de actina, promotores de inmunoglobulina, promotores de choque térmico y similares.

La transcripción se aumenta opcionalmente incluyendo una secuencia potenciadora. Los potenciadores son normalmente elementos de ADN de actuación en cis normalmente cortos, por ejemplo, de 10-500 pb, que actúan en concierto con un promotor para aumentar la transcripción. Se han aislado muchas secuencias potenciadoras de genes de mamíferos (hemoglobina, elastasa, albúmina, alfa-fetoproteína e insulina), y virus de células eucariotas. El potenciador puede cortarse y empalmarse en el vector en una posición 5' o 3' con respecto a la secuencia codificante heteróloga, pero normalmente se inserta en un sitio en 5' con respecto al promotor. Normalmente, el promotor, y si se desea, secuencias que potencian la transcripción adicionales, se eligen para optimizar la expresión en el tipo de célula huésped en el que el ADN heterólogo va a introducirse (Scharf *et al.* (1994) Heat stress promoters and transcription factors, Results Probl Cell Differ 20:125-62; Kriegler *et al.* (1990) Assembly of enhancers, promoters, and splice signals to control expression of transferred genes Methods in Enzymol 185: 512-27). Opcionalmente, el amplicón también puede contener un sitio de unión al ribosoma o un sitio interno de entrada al ribosoma (IRES) para la iniciación de la traducción.

Los vectores descritos en el presente documento también incluyen favorablemente secuencias necesarias para la terminación de la transcripción y para estabilizar el ARNm, tal como un sitio de poliadenilación o una secuencia terminadora. Tales secuencias están disponibles comúnmente de las regiones no traducidas en 5' y, ocasionalmente en 3', de ADN o ADNc virales o eucariotas. Las secuencias de señal de poliadenilación de SV40 pueden proporcionar un sitio de poliadenilación bidireccional que protege la transcripción de las moléculas de ARNm de hebra (+) de la replicación iniciada por el promotor de PolI del genoma viral de hebra (-).

Además, tal como se describió anteriormente, los vectores de expresión incluyen opcionalmente uno o más genes marcadores seleccionables para proporcionar un rasgo fenotípico para la selección de células huésped transformadas, además de los genes enumerados anteriormente, marcadores tales como dihidrofolato reductasa o resistencia a neomicina son adecuados para la selección en cultivo de células eucariotas.

El vector que contiene las secuencias de ácido nucleico apropiadas tal como se describió anteriormente, así como una secuencia de control o promotor apropiado, puede emplearse para transformar una célula huésped permitiendo la expresión de la proteína. Aunque los vectores descritos en el presente documento pueden replicarse en células bacterianas, lo más frecuentemente será deseable introducirlos en células de mamífero, por ejemplo, células Vero, células BHK, células MDCK, células 293, células COS o similares, para la finalidad de la expresión.

Tal como se describe en otra parte, las secuencias de HA y NA en el presente documento pueden estar comprendidas dentro de plásmidos implicados en reagrupamiento por rescate con plásmidos. Véanse, por ejemplo, los documentos US 2004029251 y US 2005042229. Por ejemplo, los vectores de expresión preferidos descritos en el presente documento incluyen, pero no se limitan a, vectores que comprenden el promotor de pol I y secuencias terminadoras o vectores que usan los promotores tanto de pol I como de pol II, "el sistema de promotores de polI/polII" (por ejemplo, Zobel *et al.*, Nucl. Acids Res. 1993, 21: 3607; documento US 20020164770; Neumann *et al.*, Proc. Natl. Acad. Sci. USA 1999, 96:9345; Fodor *et al.*, J. Virol. 1999, 73:9679; y documento US 20030035814). Los reagrupados producidos incluyen los genes de HA y NA dispuestos con los otros 6 genes de influenza de la cepa donadora A/Ann Arbor/6/60 (y/o derivados y modificaciones de los mismos), la estructura principal de cepa donadora

PR8, la estructura principal de cepa donadora A/Leningrad/17, etc. Se describen otras cepas de estructura principal, por ejemplo, en los documentos US 20040137013 y US 0030147916.

#### Elementos de expresión adicionales

Lo más frecuentemente, el segmento de genoma que codifica para la proteína HA y/o NA de virus influenza incluye cualquier secuencia adicional necesaria para su expresión, incluyendo traducción en una proteína viral funcional. En otras situaciones, puede emplearse un minigen, u otro constructo artificial que codifica para las proteínas virales, por ejemplo, una proteína de HA y/o NA. De nuevo, en tal caso, a menudo es deseable incluir señales de iniciación específicas que ayudan en la traducción eficaz de la secuencia codificante heteróloga. Estas señales pueden incluir, por ejemplo, el codón de iniciación ATG y secuencias adyacentes. Para garantizar la traducción de todo el inserto, el codón de iniciación se inserta en el marco de lectura correcto en relación con la proteína viral. Los elementos transcripcionales exógenos y codones de iniciación pueden ser de diversos orígenes, tanto natural como sintético. La eficacia de expresión puede potenciarse mediante la inclusión de potenciadores apropiados para el sistema celular en uso.

Si se desea, pueden incorporarse en el vector secuencias de polinucleótido que codifican para elementos expresados adicionales, tales como secuencias señal, secuencias de secreción o localización, y similares, habitualmente en marco con la secuencia de polinucleótido de interés, por ejemplo, para seleccionar como diana la expresión del polipéptido en un orgánulo, membrana o compartimento celular deseado, o para dirigir la secreción del polipéptido al espacio periplásmico o al medio de cultivo celular. Tales secuencias las conocen los expertos, e incluyen péptidos líder de secreción, secuencias de direccionamiento a orgánulos (por ejemplo, secuencias de localización nuclear, señales de retención en el RE, secuencias de tránsito mitocondrial), secuencias de anclaje/localización en la membrana (por ejemplo, secuencias de transferencia de terminación, secuencias de anclaje GPI) y similares.

Cuando se desea la traducción de un polipéptido codificado por una secuencia de ácido nucleico descrita en el presente documento, señales de iniciación específicas de traducción adicionales pueden mejorar la eficacia de traducción. Estas señales pueden incluir, por ejemplo, un codón de iniciación ATG y secuencias adyacentes, una región de IRES, etc. En algunos casos, por ejemplo, se insertan moléculas de ADNc de longitud completa o segmentos cromosómicos que incluyen una secuencia codificante que incorpora, por ejemplo, una secuencia de polinucleótido tal como se describe en el presente documento, un codón de iniciación de la traducción y elementos de secuencia asociados, en el vector de expresión apropiado simultáneamente con la secuencia de polinucleótido de interés. En tales casos, no se requieren frecuentemente señales de control de la traducción adicionales. Sin embargo, en casos en los que sólo se inserta una secuencia que codifica para el polipéptido, o una parte de la misma, a menudo se proporcionan señales de control de la traducción exógenas, incluyendo, por ejemplo, un codón de iniciación ATG, para la expresión de la secuencia relevante. El codón de iniciación se pone en el marco de lectura correcto para garantizar la transcripción de la secuencia de polinucleótido de interés. Los elementos transcripcionales exógenos y codones de iniciación pueden ser de diversos orígenes, tanto natural como sintético. La eficacia de expresión puede potenciarse mediante la inclusión de potenciadores apropiados para el sistema celular en uso (véanse, por ejemplo, ScharfD. *et al.* (1994) Results Probl Cell Differ 20:125-62; Bittner *et al.* (1987) Methods in Enzymol 153:516-544).

#### Producción de virus recombinante

Pueden modificarse por ingeniería genética virus de ARN de hebra negativa y recuperarse usando un enfoque de genética inversa recombinante (véase, por ejemplo, la patente de EE.UU. n° 5.166.057 concedida a Palese *et al.*). Tal método se aplicó originalmente para modificar por ingeniería genética genomas virales de influenza (Luytjes *et al.* (1989) Cell 59:1107-1113; Enami *et al.* (1990) Proc. Natl. Acad. Sci. USA 92:11563-11567), y se ha aplicado satisfactoriamente a una amplia variedad de virus de ARN de hebra negativa segmentados y no segmentados, por ejemplo, rabia (Schnell *et al.* (1994) EMBO J. 13: 4195-4203); VSV (Lawson *et al.* (1995) Proc. Natl. Acad. Sci. USA 92: 4477-4481); virus de las paperas (Radecke *et al.* (1995) EMBO J. 14:5773-5784); virus de la peste bovina (Baron & Barrett (1997) J. Virol. 71: 1265-1271); virus parainfluenza humano (Hoffman & Banerjee (1997) J. Virol. 71: 3272-3277; Dubin *et al.* (1997) Virology 235:323-332); SV5 (He *et al.* (1997) Virology 237:249-260); virus del moquillo canino (Gassen *et al.* (2000) J. Virol. 74:10737-44); y virus Sendai (Park *et al.* (1991) Proc. Natl. Acad. Sci. USA 88: 5537-5541; Kato *et al.* (1996) Genes to Cells 1:569-579). Los expertos en la técnica estarán familiarizados con estas y otras técnicas similares para producir virus influenza que comprenden las secuencias de HA y NA descritas en el presente documento. Los virus influenza recombinantes producidos según tales métodos son una característica de la descripción ya que son virus influenza recombinantes que comprenden uno o más ácidos nucleicos y/o polipéptidos descritos en el presente documento.

#### Cultivo celular y huéspedes de expresión

También se describen en el presente documento células huésped que se introducen (transducen, transforman o transfectan) con vectores descritos en el presente documento, y la producción de polipéptidos descritos en el presente documento mediante técnicas recombinantes. Las células huésped se modifican por ingeniería genética



(es decir, transducen, transforman o transfectan) con un vector, tal como un vector de expresión. Tal como se describió anteriormente, el vector puede estar en forma de un plásmido, una partícula viral, un fago, etc. Los ejemplos de huéspedes de expresión apropiados incluyen: células bacterianas, tales como *E. coli*, *Streptomyces* y *Salmonella typhimurium*; células fúngicas, tales como *Saccharomyces cerevisiae*, *Pichia pastoris* y *Neurospora crassa*; o células de insecto tales como *Drosophila* y *Spodoptera frugiperda*.

Lo más frecuentemente, se usan células de mamífero para cultivar las moléculas de HA y NA descritas en el presente documento. Las células huésped adecuadas para la replicación de virus influenza incluyen, por ejemplo, células Vero, células BHK, células MDCK, células 293 y células COS, incluyendo células 293T, células COS7 o similares. Comúnmente, se emplean cocultivos que incluyen dos de las líneas celulares anteriores, por ejemplo, células MDCK y o bien células 293T o bien COS a una razón, por ejemplo, de 1:1, para mejorar la eficacia de replicación. Normalmente, se cultivan las células en un medio de cultivo comercial convencional, tal como medio de Eagle modificado por Dulbecco complementado con suero (por ejemplo, suero bovino fetal al 10%), o en medio libre de suero, bajo humedad controlada y concentración de CO<sub>2</sub> adecuada para mantener el pH tamponado neutro (por ejemplo, a pH entre 7,0 y 7,2). Opcionalmente, el medio contiene antibióticos para prevenir el crecimiento bacteriano, por ejemplo, penicilina, estreptomycin, etc., y/o nutrientes adicionales, tales como L-glutamina, piruvato de sodio, aminoácidos no esenciales, complementos adicionales para promover características de crecimiento favorables, por ejemplo, tripsina,  $\beta$ -mercaptoetanol y similares.

Las células huésped modificadas por ingeniería genética pueden cultivarse en medios nutrientes convencionales modificados según sea apropiado para activar promotores, seleccionar transformantes o amplificar las secuencias de polinucleótido insertadas. Las condiciones de cultivo, tales como temperatura, pH y similares, son normalmente las usadas anteriormente con la célula huésped particular seleccionada para la expresión, y resultarán evidentes para los expertos en la técnica y en la bibliografía mencionada en el presente documento, incluyendo, por ejemplo, Freshney (1994) *Culture of Animal Cells, a Manual of Basic Technique*, 3ª edición, Wiley-Liss, Nueva York y la bibliografía mencionada en la misma. Otra bibliografía útil incluye, por ejemplo, Paul (1975) *Cell and Tissue Culture*, 5ª ed., Livingston, Edinburg; Adams (1980) *Laboratory Techniques in Biochemistry and Molecular Biology-Cell Culture for Biochemists*, Work y Burdon (eds.) Elsevier, Amsterdam. Los detalles adicionales referentes a procedimientos de cultivo tisular de particular interés en la producción de virus influenza *in vitro* incluyen, por ejemplo, Merten *et al.* (1996) *Production of influenza virus in cell cultures for vaccine preparation*. En Cohen y Shafferman (eds.) *Novel Strategies in Design and Production of Vaccines*. Adicionalmente, se determinan fácilmente variaciones en tales procedimientos adaptados a la presente descripción a través de experimentación rutinaria y serán familiares para los expertos en la técnica.

Pueden cultivarse células para la producción de virus influenza (por ejemplo, que tiene las secuencias de HA y/o NA de la descripción) en medio libre de suero o que contiene suero. En algunos casos, por ejemplo, para la preparación de virus purificados, normalmente es deseable hacer crecer las células huésped en condiciones libres de suero. Las células pueden cultivarse a pequeña escala, por ejemplo, menos de 25 ml de medio, tubos de cultivo o matraces o en matraces grandes con agitación, en frascos rotatorios, o en perlas microportadoras (por ejemplo, perlas microportadoras de DEAE-dextrano, tales como Dormacell, Pfeifer & Langen; Superbead, Flow Laboratories; perlas de copolímero de estireno-trimetilamina, tales como Hillex, SoloHill, Ann Arbor) en matraces, frascos o cultivos en reactor. Las perlas microportadoras son pequeñas esferas (en el intervalo de 100-200 micrómetros de diámetro) que proporcionan una gran área superficial para crecimiento celular adherente por volumen de cultivo celular. Por ejemplo, un único litro de medio puede incluir más de 20 millones de perlas microportadoras que proporcionan más de 8000 centímetros cuadrados de superficie de crecimiento. Para la producción comercial de virus, por ejemplo, para la producción de vacunas, a menudo es deseable cultivar las células en un biorreactor o fermentador. Están disponibles biorreactores en volúmenes de desde menos de 1 litro hasta en exceso de 100 litros, por ejemplo, biorreactor Cyto3 (Osmonics, Minnetonka, MN); biorreactores NBS (New Brunswick Scientific, Edison, NJ); biorreactores a escala de laboratorio y comercial de B. Braun Biotech International (B. Braun Biotech, Melsungen, Alemania).

Independientemente del volumen de cultivo, en muchos aspectos deseados de la presente descripción, es importante que los cultivos se mantengan a una temperatura apropiada, para garantizar una recuperación eficaz de virus influenza recombinante y/o reagrupado usando sistemas de múltiples plásmidos dependientes de temperatura (véase, por ejemplo, Sistema de múltiples plásmidos para la producción de virus influenza, documento US 2004029251), calentamiento de disoluciones de virus para la filtración, etc. Normalmente, se emplea un regulador, por ejemplo, un termostato, u otro dispositivo para detectar y mantener la temperatura del sistema de cultivo celular y/u otra disolución, para garantizar que la temperatura está al nivel correcto durante el periodo apropiado (por ejemplo, replicación del virus, etc.).

En la descripción en el presente documento (por ejemplo, cuando van a producirse virus reagrupados a partir de segmentos en vectores), se introducen (por ejemplo, transfectan) vectores que comprenden segmentos de genoma de influenza en células huésped según métodos bien conocidos en la técnica para introducir ácidos nucleicos heterólogos en células eucariotas, incluyendo, por ejemplo, coprecipitación con fosfato de calcio, electroporación, microinyección, lipofección y transfección empleando reactivos de transfección de poliamina. Por ejemplo, pueden transfectarse vectores, por ejemplo, plásmidos en células huésped, tales como células COS, células 293T o

combinaciones de células COS o 293T y células MDCK, usando el reactivo de transfección de poliamina TransIT-LT1 (Mirus) según las instrucciones del fabricante con el fin de producir virus reagrupados, etc. Por tanto, en un ejemplo, se introduce aproximadamente 1 µg de cada vector en una población de células huésped con aproximadamente 2 µl de TransIT-LT1 diluidos en 160 µl de medio, preferiblemente medio libre de suero, en un volumen total de 200 µl. Se incuban las mezclas de ADN:reactivo de transfección a temperatura ambiente durante 45 minutos seguido por adición de 800 µl de medio. Se añade la mezcla de transfección a las células huésped, y se cultivan las células tal como se describe mediante otros métodos bien conocidos por los expertos en la técnica. Por consiguiente, para la producción de virus recombinantes o reagrupados en cultivo celular, se mezclan vectores que incorporan cada uno de los 8 segmentos de genoma (PB2, PB1, PA, NP, M, NS, HA y NA) con aproximadamente 20 µl de TransIT-LT1 y se transfectan en células huésped. Opcionalmente, se sustituye el medio que contiene suero antes de la transfección por medio libre de suero, por ejemplo, Opti-MEM I, y se incuba durante 4-6 horas.

Alternativamente, puede emplearse electroporación para introducir tales vectores que incorporan segmentos de genoma de influenza en células huésped. Por ejemplo, vectores de plásmido que incorporan un virus influenza A o influenza B se introducen favorablemente en células Vero usando electroporación según el siguiente procedimiento. En resumen, se resuspenden aproximadamente  $5 \times 10^6$  células Vero, por ejemplo, hechas crecer en medio de Eagle modificado (MEM) complementado con suero bovino fetal al 10% (FBS), en 0,4 ml de OptiMEM y se colocan en una cubeta de electroporación. Se añaden veinte microgramos de ADN en un volumen de hasta 25 µl a las células en la cubeta, que entonces se mezclan suavemente dando golpecitos. Se realiza la electroporación según las instrucciones del fabricante (por ejemplo, BioRad Gene Pulser II con Capacitance Extender Plus conectado) a 300 voltios, 950 microfaradios con un tiempo constante de entre 28-33 ms. Las células vuelven a mezclarse dando golpecitos suaves y aproximadamente 1-2 minutos tras la electroporación se añaden 0,7 ml de MEM con FBS al 10% directamente a la cubeta. Entonces se transfieren las células a dos pocillos de una placa de cultivo tisular de 6 pocillos convencional que contiene 2 ml de MEM, FBS al 10%. Se lava la cubeta para recuperar cualquier célula restante y se divide la suspensión de lavado entre los dos pocillos. El volumen final es de aproximadamente 3,5 ml. Después, se incuban las células en condiciones permisivas para el crecimiento viral, por ejemplo, a aproximadamente 33°C para cepas adaptadas al frío.

En células huésped de mamífero, pueden utilizarse varios sistemas de expresión, tales como sistemas a base de virus. En casos en los que se usa un adenovirus como vector de expresión, se liga opcionalmente una secuencia codificante en un complejo de transcripción/traducción de adenovirus que consiste en el promotor tardío y la secuencia líder tripartita. La inserción en una región E1 o E3 no esencial del genoma viral dará como resultado un virus viable que puede expresar los polipéptidos de interés en células huésped infectadas (Logan y Shenk (1984) Proc Natl Acad Sci 81:3655-3659). Además, pueden usarse potenciadores de la transcripción, tales como el potenciador del virus del sarcoma de Rous (VSR), para aumentar la expresión en células huésped de mamífero.

Se elige opcionalmente una cepa de célula huésped por su capacidad para modular la expresión de las secuencias insertadas o para procesar la proteína expresada del modo deseado. Tales modificaciones de la proteína incluyen, pero no se limitan a, acetilación, carboxilación, glicosilación, fosforilación, lipidación y acilación. El procesamiento postraducciona de la proteína, que escinde una forma precursora para dar una forma madura, es algunas veces importante para corregir la inserción, el plegamiento y/o la función. Adicionalmente, también es importante la ubicación apropiada dentro de una célula huésped (por ejemplo, en la superficie de la célula). Diferentes células huésped tales como COS, CHO, BHK, MDCK, 293, 293T, COS7, etc. tienen maquinaria celular específica y mecanismos característicos para tales actividades postraduccionales y pueden elegirse para garantizar el procesamiento y la modificación correctos de la presente proteína foránea introducida.

Para la producción a largo plazo, de alto rendimiento de proteínas recombinantes codificadas por, o que tienen subsecuencias codificadas por, los polinucleótidos descritos en el presente documento, se usan opcionalmente sistemas de expresión estable. Por ejemplo, se transfectan líneas celulares, que expresan de manera estable un polipéptido descrito en el presente documento, usando vectores de expresión que contienen orígenes de replicación virales o elementos de expresión endógenos y un gen marcador seleccionable. Por ejemplo, tras la introducción del vector, se deja que las células crezcan durante 1-2 días en un medio enriquecido antes de que se cambien a medios selectivos. El fin del marcador seleccionable es conferir resistencia a la selección, y su presencia permite el crecimiento y la recuperación de células que expresan satisfactoriamente las secuencias introducidas. Por tanto, pueden proliferar grupos resistentes de células transformadas de manera estable, por ejemplo, derivadas de un único tipo de célula, usando técnicas de cultivo tisular apropiadas para el tipo de célula.

Las células huésped transformadas con una secuencia de nucleótidos que codifica para un polipéptido descrito en el presente documento se cultivan opcionalmente en condiciones adecuadas para la expresión y recuperación de la proteína codificada del cultivo celular. Las células que expresan dicha proteína pueden clasificarse, aislarse y/o purificarse. La proteína o fragmento de la misma producido por una célula recombinante puede secretarse, unirse a la membrana o retenerse intracelularmente, dependiendo de la secuencia (por ejemplo, dependiendo de proteínas de fusión que codifican para una señal de retención en la membrana o similar) y/o el vector usado.

Los productos de expresión que corresponden a los ácidos nucleicos descritos en el presente documento también

pueden producirse en células no animales tales como plantas, levaduras, hongos, bacterias y similares. Además de Sambrook, Berger y Ausubel, todos citados a continuación, pueden encontrarse detalles referentes al cultivo celular en Payne *et al.* (1992) *Plant Cell and Tissue Culture in Liquid Systems* John Wiley & Sons, Inc. Nueva York, NY; Gamborg y Phillips (eds.) (1995) *Plant Cell, Tissue and Organ Culture; Fundamental Methods* Springer Lab Manual, Springer-Verlag (Berlín Heidelberg Nueva York) y Atlas y Parks (eds.) *The Handbook of Microbiological Media* (1993) CRC Press, Boca Raton, FL.

En sistemas bacterianos, pueden seleccionarse varios vectores de expresión dependiendo del uso previsto para el producto expresado. Por ejemplo, cuando se necesitan grandes cantidades de un polipéptido o fragmentos del mismo para la producción de anticuerpos, se emplean favorablemente vectores que dirigen la expresión a alto nivel de proteínas de fusión que se purifican fácilmente. Tales vectores incluyen, pero no se limitan a, vectores de expresión y clonación de *E. coli* multifuncionales tales como BLUESCRIPT (Stratagene), en los que la secuencia codificante de interés, por ejemplo, secuencias que comprenden las encontradas en el presente documento, etc., puede ligarse en el vector en marco con secuencias para la traducción amino terminal que se inicia con metionina y los 7 residuos posteriores de beta-galactosidasa que producen una proteína de fusión de beta-galactosidasa catalíticamente activa; vectores pIN (Van Heeke & Schuster (1989) *J Biol Chem* 264:5503-5509); vectores pET (Novagen, Madison WI); y similares. De manera similar, en la levadura *Saccharomyces cerevisiae*, pueden usarse varios vectores que contienen promotores constitutivos o inducibles, tales como factor alfa, alcohol oxidasa y PGH para la producción de los productos de expresión deseados. Para revisiones, véase Ausubel, citado a continuación, y Grant *et al.*, (1987); *Methods in Enzymology* 153:516-544.

#### Hibridación de ácidos nucleicos

Puede usarse hibridación comparativa para identificar ácidos nucleicos (por ejemplo, SEC ID N° 5-6) incluyendo variaciones conservadoras de ácidos nucleicos descritos en el presente documento. Este método de hibridación comparativa es un método preferido de distinción de ácidos nucleicos. Además, ácidos nucleicos diana que hibridan con los ácidos nucleicos representados por los mostrados en el presente documento en condiciones de alta, ultra alta y ultra-ultra alta rigurosidad, son características de la descripción. Los ejemplos de tales ácidos nucleicos incluyen aquellos con una o unas cuantas sustituciones de ácido nucleico silenciosas o conservadoras en comparación con una secuencia de ácido nucleico dada.

Se dice que un ácido nucleico diana de prueba hibrida específicamente con un ácido nucleico sonda cuando hibrida al menos la mitad de bien con la sonda que con la diana complementaria perfectamente apareada, es decir, con una razón señal - ruido al menos la mitad de alta que la hibridación de la sonda y la diana en condiciones en las que una sonda perfectamente apareada se une a una diana complementaria perfectamente apareada con una razón señal - ruido que es al menos aproximadamente 5x-10x tan alta como la observada para la hibridación con cualquiera de los ácidos nucleicos no apareados.

Los ácidos nucleicos "hibridan" cuando se asocian, normalmente en disolución. Los ácidos nucleicos hibridan debido a una variedad de fuerzas fisicoquímicas bien caracterizadas, tales como puentes de hidrógeno, exclusión de disolvente, apilamiento de bases y similares. Se conocen bien en la técnica numerosos protocolos para la hibridación de ácidos nucleicos. Se encuentra una guía extensa para la hibridación de ácidos nucleicos en Tijssen (1993) *Laboratory Techniques in Biochemistry and Molecular Biology—Hybridization with Nucleic Acid Probes* parte I capítulo 2, "Overview of principles of hybridization and the strategy of nucleic acid probe assays," (Elsevier, Nueva York), así como en Ausubel, Sambrook, y Berger y Kimmel, todos citados a continuación. Hames y Higgins (1995) *Gene Probes 1* IRL Press en Oxford University Press, Oxford, Inglaterra, (Hames y Higgins 1) y Hames y Higgins (1995) *Gene Probes 2* IRL Press en Oxford University Press, Oxford, Inglaterra (Hames y Higgins 2) proporcionan detalles sobre la síntesis, el marcaje, la detección y cuantificación de ADN y ARN, incluyendo oligonucleótidos.

Un ejemplo de condiciones de hibridación rigurosas para la hibridación de ácidos nucleicos complementarios que tienen más de 100 residuos complementarios sobre un filtro en una transferencia de tipo Southern o Northern es un 50% de formalina con 1 mg de heparina a 42°C, llevándose a cabo la hibridación durante la noche. Un ejemplo de condiciones de lavado rigurosas comprende un lavado con 0,2x SSC a 65°C durante 15 minutos (véase, Sambrook, citado a continuación, para una descripción del tampón de SSC y otros parámetros de hibridación de ácidos nucleicos). A menudo, al lavado de alta rigurosidad le precede un lavado de baja rigurosidad para eliminar la señal de sonda de fondo. Un ejemplo de lavado de baja rigurosidad es 2x SSC a 40°C durante 15 minutos. En general, una razón señal - ruido de 5x (o superior) a la observada para una sonda no relacionada en el ensayo de hibridación particular indica detección de una hibridación específica.

Tras la hibridación, pueden eliminarse los ácidos nucleicos no hibridados mediante una serie de lavados, cuya rigurosidad puede ajustarse dependiendo de los resultados deseados. Las condiciones de lavado de baja rigurosidad (por ejemplo, usando concentración de sal superior y temperatura inferior) aumentan la sensibilidad, pero pueden producir señales de hibridación no específicas y señales de fondo altas. Las condiciones de rigurosidad superior (por ejemplo, usando concentración de sal inferior y temperatura superior que está más cerca de la T<sub>m</sub>) disminuyen la señal de fondo, normalmente con la señal específica restante específicamente. Véase, también, Rapley, R. y Walker, J.M. eds., *Molecular Biomethods Handbook* (Humana Press, Inc. 1998), "Condiciones de lavado

de hibridación rigurosa" en el contexto de experimentos de hibridación de ácidos nucleicos tales como hibridaciones de tipo Southern y Northern dependen de la secuencia, y son diferentes bajo parámetros medioambientales diferentes. Se encuentra una guía extensa para la hibridación de ácidos nucleicos en Tijssen (1993), citado anteriormente, y en Hames y Higgins, 1 y 2. Las condiciones de lavado e hibridación rigurosas pueden determinarse fácilmente de manera empírica para cualquier ácido nucleico de prueba. Por ejemplo, al determinar las condiciones de lavado e hibridación altamente rigurosas, las condiciones de lavado e hibridación se aumentan gradualmente (por ejemplo, aumentando la temperatura, disminuyendo la concentración de sal, aumentando la concentración de detergente y/o aumentando la concentración de disolventes orgánicos tales como formalina en la hibridación y el lavado), hasta que se cumple un conjunto seleccionado de criterios. Por ejemplo, las condiciones de lavado e hibridación se aumentan gradualmente hasta que una sonda se une a una diana complementaria perfectamente apareada con una razón señal - ruido que es al menos 5x tan alta como la observada para la hibridación de la sonda con una diana no apareada.

En general, una razón señal - ruido de al menos 2x (o superior, por ejemplo, al menos 5x, 10x, 20x, 50x, 100x, o más) la observada para una sonda no relacionada en el ensayo de hibridación particular indica detección de una hibridación específica. La detección de hibridación al menos rigurosa entre dos secuencias en el contexto de la presente descripción indica similitud estructural relativamente fuerte con, por ejemplo, los ácidos nucleicos proporcionados en las listas de secuencias en el presente documento.

Se seleccionan condiciones "muy rigurosas" para que sean iguales al punto de fusión térmica ( $T_m$ ) para una sonda particular. La  $T_m$  es la temperatura (a pH y fuerza iónica definidos) a la que el 50% de la secuencia de prueba hibrida con una sonda perfectamente apareada. Para los fines de la presente descripción, generalmente, se seleccionan condiciones de hibridación y lavado "altamente rigurosas" para que sean aproximadamente 5°C inferiores a la  $T_m$  para la secuencia específica a un pH y fuerza iónica definidos (tal como se indica a continuación, también puede hacerse referencia a condiciones altamente rigurosas en términos comparativos). Pueden identificarse secuencias diana que están estrechamente relacionadas o son idénticas a la secuencia de nucleótidos de interés (por ejemplo, "sonda") en condiciones rigurosas o altamente rigurosas. Condiciones de rigurosidad inferior son apropiadas para secuencias que son menos complementarias.

Condiciones de hibridación y lavado de "rigurosidad ultraalta" son aquéllas en las que la rigurosidad de las condiciones de hibridación y lavado se aumentan hasta que la razón señal - ruido para la unión de una sonda a un ácido nucleico diana complementario es al menos 10x tan alta como la observada para la hibridación con cualquier ácido nucleico diana no apareado. Un ácido nucleico diana que hibrida con una sonda en tales condiciones, con una razón señal - ruido de al menos la mitad de la del ácido nucleico diana complementario perfectamente apareado se dice que se une a la sonda en condiciones de rigurosidad ultraalta.

Al determinar las condiciones de hibridación y lavado rigurosas o altamente rigurosas (o incluso hibridación más rigurosa), las condiciones de hibridación y lavado se aumentan gradualmente (por ejemplo, aumentando la temperatura, disminuyendo la concentración de sal, aumentando la concentración de detergente y/o aumentando la concentración de disolventes orgánicos, tales como formamida, en la hibridación o el lavado), hasta que se cumple un conjunto seleccionado de criterios. Por ejemplo, las condiciones de hibridación y lavado se aumentan gradualmente hasta que una sonda que comprende una o más secuencias de polinucleótido de la descripción, por ejemplo, secuencias o subsecuencias únicas seleccionadas de las facilitadas en el presente documento (por ejemplo, SEC ID N° 1-2) y/o secuencias de polinucleótido complementarias, se unen a una diana complementaria perfectamente apareada (de nuevo, un ácido nucleico que comprende una o más secuencias o subsecuencias de ácido nucleico seleccionadas de las facilitadas en el presente documento y/o secuencias de polinucleótido complementarias de las mismas), con una razón señal - ruido que es al menos 2x (y opcionalmente 5x, 10x o 100x o más) tan alta como la observada para la hibridación de la sonda con una diana no apareada (por ejemplo, una secuencia de polinucleótido que comprende una o más secuencias o subsecuencias seleccionadas de secuencias de influenza conocidas presentes en bases de datos públicas tales como GenBank en el momento de la presentación, y/o secuencias de polinucleótido complementarias de las mismas), según se desee.

Usando los polinucleótidos descritos en el presente documento, o subsecuencias de los mismos, pueden obtenerse ácidos nucleicos diana novedosos. Por ejemplo, tales ácidos nucleicos diana incluyen secuencias que hibridan en condiciones rigurosas con una sonda oligonucleotídica única correspondiente a cualquiera de los polinucleótidos de SEC ID N° 1-2).

De manera similar, pueden determinarse niveles de rigurosidad incluso superiores aumentando gradualmente las condiciones de hibridación y/o lavado del ensayo de hibridación relevante. Por ejemplo, aquellos en los que la rigurosidad de las condiciones de hibridación y lavado se aumentan hasta que la razón señal - ruido para la unión de la sonda al ácido nucleico diana complementario perfectamente apareado es al menos 10X, 20X, 50X, 100X o 500X o más tan alta como la observada para la hibridación con cualquier ácido nucleico diana no apareado. La señal particular dependerá del marcador usado en el ensayo relevante, por ejemplo, un marcador fluorescente, un marcador colorimétrico, un marcador radiactivo o similar. Un ácido nucleico diana que hibrida con una sonda en tales condiciones, con una razón señal - ruido de al menos la mitad de la del ácido nucleico diana complementario perfectamente apareado, se dice que se une a la sonda en condiciones de rigurosidad ultra-ultraalta.

Los ácidos nucleicos que no hibridan entre sí en condiciones rigurosas son todavía sustancialmente idénticos si los polipéptidos que codifican son sustancialmente idénticos. Esto se produce, por ejemplo, cuando se crea una copia de un ácido nucleico usando la degeneración de codones máxima permitida por el código genético.

#### Clonación, mutagénesis y expresión de biomoléculas de interés

Textos generales que describen técnicas de biología molecular, que son aplicables a la presente invención, tales como clonación, mutación, cultivo celular y similares, incluyen Berger y Kimmel, Guide to Molecular Cloning Techniques, Methods in Enzymology, volumen 152 Academic Press, Inc., San Diego, CA (Berger); Sambrook *et al.*, Molecular Cloning - A Laboratory Manual (3ª Ed.), Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, Nueva York, 2000 ("Sambrook") y Current Protocols in Molecular Biology, F. M. Ausubel *et al.*, eds., Current Protocols, una empresa conjunta entre Greene Publishing Associates, Inc. y John Wiley & Sons, Inc., (complementadas hasta 2002) ("Ausubel"). Estos textos describen la mutagénesis, el uso de vectores, promotores y muchos otros temas relevantes relacionados con, por ejemplo, la generación de moléculas de HA y/o NA, etc.

También se describen en el presente documento diversos tipos de mutagénesis, por ejemplo, para producir y/o aislar, por ejemplo, moléculas de HA y/o NA novedosas o recién aisladas y/o para modificar/mutar adicionalmente los polipéptidos (por ejemplo, moléculas de HA y/o NA como en la SEC ID N° 11-12) descrita en el presente documento. Incluyen, pero no se limitan, mutagénesis puntual aleatoria, dirigida al sitio, recombinación homóloga (intercambio de ADN), mutagénesis usando moldes que contienen uracilo, mutagénesis dirigida a oligonucleótidos, mutagénesis de ADN modificada por fosforotioato, mutagénesis usando ADN bicatenario incompleto, o similares. Métodos adecuados adicionales incluyen reparación de apareamientos erróneos puntuales, mutagénesis usando cadenas de huésped con reparación deficiente, restricción-selección y restricción-purificación, mutagénesis por delección, mutagénesis mediante síntesis génica total, reparación de rotura de cadena doble y similares. También se describe en el presente documento la mutagénesis, por ejemplo, que implica constructos quiméricos. La mutagénesis puede guiarse mediante información conocida de la molécula que se produce de manera natural o la molécula que se produce de manera natural alterada o mutada, por ejemplo, mediante comparaciones de secuencia, propiedades físicas, estructura cristalina o similares.

Los textos anteriores y los ejemplos encontrados en el presente documento describen estos procedimientos, al igual que las siguientes publicaciones (y bibliografía citada dentro de ellas): Sieber, *et al.*, Nature Biotechnology, 19:456-460 (2001); Ling *et al.*, Approaches to DNA mutagenesis: an overview, Anal Biochem 254(2): 157-178 (1997); Dale *et al.*, Oligonucleotide-directed random mutagenesis using the phosphorothioate method, Methods Mol Biol 57:369-374 (1996); 1. A. Lorimer, I. Pastan, Nucleic Acid Res 23, 3067-8 (1995); W. P. C. Stemmer, Nature 370, 389-91 (1994); Arnold, Protein engineering for unusual environments, Current Opinion in Biotechnology 4:450-455 (1993); Bass *et al.*, Mutant Trp repressors with new DNA-binding specificities, Science 242:240-245 (1988); Fritz *et al.*, Oligonucleotide-directed construction of mutations: a gapped duplex DNA procedure without enzymatic reactions *in vitro*, Nucl Acids Res 16: 6987-6999 (1988); Kramer *et al.*, Improved enzymatic *in vitro* reactions in the gapped duplex DNA approach to oligonucleotide-directed construction of mutations, Nucl Acids Res 16: 7207 (1988); Sakamar y Khorana, Total synthesis and expression of a gene for the  $\alpha$ -subunit of bovine rod outer segment guanine nucleotide-binding protein (transducin), Nucl Acids Res 14: 6361-6372 (1988); Sayers *et al.*, Y-T Exonucleases in phosphorothioate-based oligonucleotide-directed mutagenesis, Nucl Acids Res 16:791-802 (1988); Sayers *et al.*, Strand specific cleavage of phosphorothioate-containing DNA by reaction with restriction endonucleases in presence of ethidium bromide, (1988) Nucl Acids Res 16: 803-814; Carter, Improved oligonucleotide-directed mutagenesis using M13 vectors, Methods in Enzymol 154: 382-403 (1987); Kramer & Fritz Oligonucleotide-directed construction of mutations via gapped duplex DNA, Methods in Enzymol 154:350-367 (1987); Kunkel, The efficiency of oligonucleotide directed mutagenesis, en Nucleic Acids & Molecular Biology (Eckstein, F. y Lilley, D.M.J. eds., Springer Verlag, Berlín) (1987); Kunkel *et al.*, Rapid and efficient site-specific mutagenesis without phenotypic selection, Methods in Enzymol 154, 367-382 (1987); Zoller & Smith, Oligonucleotide-directed mutagenesis: a simple method using two oligonucleotide primers and a single-stranded DNA template, Methods in Enzymol 154: 329-350 (1987); Carter, Site-directed mutagenesis, Biochem J 237:1-7 (1986); Eghtedarzadeh & Henikoff, Use of oligonucleotides to generate large deletions, Nucl Acids Res 14: 5115 (1986); Mandecki, Oligonucleotide-directed double-strand break repair in plasmids of *Escherichia coli*: a method for site-specific mutagenesis, Proc Natl Acad Sci USA, 83: 7177-7181 (1986); Nakamaye & Eckstein, Inhibition of restriction endonuclease Nci I cleavage by phosphorothioate groups and its application to oligonucleotide-directed mutagenesis, Nucl Acids Res 14: 9679-9698 (1986); Wells *et al.*, Importance of hydrogen-bond formation in stabilizing the transition state of subtilisin, Phil Trans R Soc Lond A 317: 415-423 (1986); Botstein & Shortle, Strategies and applications of *in vitro* mutagenesis, Science 229:1193-1201(1985); Carter *et al.*, Improved oligonucleotide site-directed mutagenesis using M13 vectors, Nucl Acids Res 13: 4431-4443 (1985); Grundström *et al.*, Oligonucleotide-directed mutagenesis by microscale 'shot-gun' gene synthesis, Nucl Acids Res 13: 3305-3316 (1985); Kunkel, Rapid and efficient site-specific mutagenesis without phenotypic selection, Proc Natl Acad Sci USA 82:488-492 (1985); Smith, *In vitro* mutagenesis, Ann Rev Genet 19:423-462(1985); Taylor *et al.*, The use of phosphorothioate-modified DNA in restriction enzyme reactions to prepare nicked DNA, Nucl Acids Res 13: 8749-8764 (1985); Taylor *et al.*, The rapid generation of oligonucleotide-directed mutations at high frequency using phosphorothioate-modified DNA, Nucl Acids Res 13: 8765-8787 (1985); Wells *et al.*, Cassette mutagenesis: an efficient method for generation of multiple mutations at defined sites, Gene

34:315-323 (1985); Kramer *et al.*, The gapped duplex DNA approach to oligonucleotide-directed mutation construction, Nucl Acids Res 12: 9441-9456 (1984); Kramer *et al.*, Point Mismatch Repair, Cell 38:879-887 (1984); Nambiar *et al.*, Total synthesis and cloning of a gene coding for the ribonuclease S protein, Science 223: 1299-1301 (1984); Zoller & Smith, Oligonucleotide-directed mutagenesis of DNA fragments cloned into M13 vectors, Methods in Enzymol 100:468-500 (1983); y Zoller & Smith, Oligonucleotide-directed mutagenesis using M13-derived vectors: an efficient and general procedure for the production of point mutations in any DNA fragment, Nucl Acids Res 10:6487-6500 (1982). Detalles adicionales sobre muchos de los métodos anteriores pueden encontrarse en Methods in Enzymol Volumen 154, que también describe controles útiles para problemas de corrección de errores con diversos métodos de mutagénesis, aislamiento de genes, expresión y otros.

Los oligonucleótidos, por ejemplo, para su uso en mutagénesis por ejemplo mutando bibliotecas de moléculas de HA y/o NA o alterándolas, normalmente se sintetizan químicamente según el método de triéster de fosforamidita en fase sólida descrito por Beaucage y Caruthers, Tetrahedron Letts 22(20):1859-1862, (1981) por ejemplo usando un sintetizador automatizado, tal como se describe en Needham-VanDevanter *et al.*, Nucleic Acid Res, 12:6159-6168 (1984).

Además, esencialmente puede adaptarse o encargarse de manera convencional cualquier ácido nucleico a partir de cualquiera de una variedad de fuentes comerciales, tales como The Midland Certified Reagent Company (mcrco@oligos.com), The Great American Gene Company (www.genco.com), ExpressGen Inc. (www.expressgen.com), Operon Technologies Inc. (Alameda, CA) y muchas otras. De manera similar, también pueden encargarse a medida péptidos y anticuerpos a partir de cualquiera de una variedad de fuentes, tales como PeptidoGenic (disponible en pkim@ccnet.com), HTI Bio-products, Inc. (www.htibio.com), BMA Biomedicals Ltd. (R.U.), Bio.Synthesis, Inc., y muchas otras.

También se describen en el presente documento células huésped y organismos que comprenden una molécula de HA y/o NA u otro polipéptido y/o ácido nucleico de por ejemplo, SEC ID N° 1-2. Las células huésped se modifican mediante ingeniería genética (por ejemplo, se transforman, transducen o transfectan) con los vectores descritos en el presente documento, que pueden ser, por ejemplo, un vector de clonación o un vector de expresión. El vector puede estar, por ejemplo, en forma de un plásmido, una bacteria, un virus, un polinucleótido desnudo o un polinucleótido conjugado. Los vectores se introducen en células y/o microorganismos mediante métodos convencionales que incluyen electroporación (véase, From *et al.*, Proc Natl Acad Sci USA 82, 5824 (1985)), infección mediante vectores virales, penetración balística a alta velocidad mediante partículas pequeñas con el ácido nucleico o bien dentro de la matriz de perlas o partículas pequeñas, o bien sobre la superficie (Klein *et al.*, Nature 327, 70-73 (1987)). Berger, Sambrook y Ausubel proporcionan una variedad de métodos de transformación adecuados. Véase anteriormente.

Están disponibles varios métodos bien conocidos de introducción de ácidos nucleicos diana en células bacterianas, pudiendo usarse cualquiera de ellos en la presente invención. Éstos incluyen: fusión de las células receptoras con protoplastos bacterianos que contienen el ADN, electroporación, bombardeo con proyectiles e infección con vectores virales, etc. Pueden usarse células bacterianas para amplificar el número de plásmidos que contienen construcciones de ADN de la presente invención. Las bacterias se hacen crecer hasta fase logarítmica y pueden aislarse los plásmidos dentro de las bacterias mediante una variedad de métodos conocidos en la técnica (véase, por ejemplo, Sambrook). Además, están disponibles comercialmente una multitud de kits para la purificación de plásmidos procedentes de bacterias (véase, por ejemplo, EasyPrep™, FlexiPrep™, ambos de Pharmacia Biotech; StrataClean™, de Stratagene; y, QIAprep™ de Qiagen). Los plásmidos aislados y purificados se manipulan entonces adicionalmente para producir otros plásmidos, se usan para transfectar células o se incorporan en vectores relacionados para infectar organismos. Los vectores típicos contienen secuencias terminadoras de transcripción y traducción, secuencias de inicio de transcripción y traducción, y promotores útiles para la regulación de la expresión del ácido nucleico diana particular. Los vectores comprenden opcionalmente casetes de expresión genéricos que contienen al menos una secuencia terminadora independiente, secuencias que permiten la replicación del casete en eucariotas, o procariotas, o ambos, (por ejemplo, vectores lanzadera) y marcadores de selección para sistemas tanto procariotas como eucariotas. Los vectores son adecuados para la replicación y la integración en procariotas, eucariotas, u opcionalmente ambos. Véanse, Gilman & Smith, Gene 8:81 (1979); Roberts, *et al.*, Nature, 328:731 (1987); Schneider, B., *et al.*, Protein Expr Purif 6435:10 (1995); Ausubel, Sambrook, Berger (todos citados anteriormente). Se proporciona un catálogo de bacterias y bacteriófagos útiles para la clonación, por ejemplo, por la ATCC, por ejemplo, The ATCC Catalogue of Bacteria and Bacteriophages (1992) Ghema *et al.* (eds.) publicado por la ATCC. Procedimientos básicos adicionales para secuenciación, clonación y otros aspectos de biología molecular y consideraciones teóricas subyacentes también se encuentran en Watson *et al.* (1992) Recombinant DNA, Segunda edición, Scientific American Books, NY. Véase, anteriormente. Vectores adicionales útiles con las secuencias en el presente documento se ilustran anteriormente en la sección relativa a la producción de virus influenza para vacunas y la bibliografía citada en ella.

#### PRODUCCIÓN Y RECUPERACIÓN DE POLIPÉPTIDOS

Tras la transducción de una cepa o línea celular huésped adecuada y el crecimiento de las células huésped hasta una densidad celular apropiada, se induce el promotor seleccionado mediante medios apropiados (por ejemplo,

cambio de temperatura o inducción química) y las células se cultivan durante un periodo adicional. En algunas realizaciones, se recupera entonces un producto polipeptídico secretado, por ejemplo, un polipéptido de HA y/o NA en forma de proteína de fusión secretada, etc., del medio de cultivo. En otras realizaciones, se produce una partícula de virus de la invención que contiene un polipéptido de HA y/o NA procedente de la célula. Alternativamente, las células pueden recogerse mediante centrifugación, alterarse mediante medios físicos o químicos, y el extracto bruto resultante puede retenerse para su purificación adicional. Las células eucariotas o microbianas empleadas en la expresión de proteínas pueden alterarse mediante cualquier método conveniente, incluyendo ciclos de congelación-descongelación, sonicación, alteración mecánica, o uso de agentes de lisis celular, u otros métodos que conocen bien los expertos en la técnica. Adicionalmente, pueden utilizarse células que expresan un producto polipeptídico de HA y/o NA descrito en el presente documento sin separar el polipéptido de la célula. En tales situaciones, el polipéptido se expresa opcionalmente sobre la superficie celular y se examina por tanto (por ejemplo, teniendo moléculas de HA y/o NA por ejemplo, que comprenden proteínas de fusión o similares) sobre la superficie celular, etc.

Los polipéptidos expresados pueden recuperarse y purificarse a partir de cultivos celulares recombinantes mediante cualquiera de varios métodos bien conocidos en la técnica, incluyendo precipitación con etanol o sulfato de amonio, extracción con ácido, cromatografía de intercambio aniónico o catiónico, cromatografía en fosfocelulosa, cromatografía de interacción hidrófoba, cromatografía de afinidad (por ejemplo, usando cualquiera de los sistemas de marcaje conocidos por los expertos en la técnica), cromatografía en hidroxilapatita y cromatografía en lectina. Pueden usarse etapas de plegamiento de proteínas, según se desee, para completar la configuración de la proteína madura. Además, puede emplearse cromatografía de líquidos de alta resolución (HPLC) en las etapas de purificación finales. Además de la bibliografía mencionada en el presente documento, se conocen bien en la técnica una variedad de métodos de purificación, incluyendo, por ejemplo, los explicados en Sandana (1997) *Bioseparation of Proteins*, Academic Press, Inc.; y Bollag *et al.* (1996) *Protein Methods*, 2ª Edición Wiley-Liss, NY; Walker (1996) *The Protein Protocols Handbook* Humana Press, NJ, Harris y Angal (1990) *Protein Purification Applications: A Practical Approach* IRL Press at Oxford, Oxford, Inglaterra; Harris y Angal, *Protein Purification Methods: A Practical Approach* IRL Press at Oxford, Oxford, Inglaterra; Scopes (1993) *Protein Purification: Principles and Practice* 3ª Edición Springer Verlag, NY; Janson y Ryden (1998) *Protein Purification: Principles, High Resolution Methods and Applications*. Segunda edición, Wiley-VCH, NY; y Walker (1998) *Protein Protocols on CD-ROM* Humana Press, NJ.

Cuando los polipéptidos expresados se producen en virus, los virus se recuperan normalmente del medio de cultivo, en el que se han hecho crecer las células infectadas (transfectadas). Normalmente, el medio bruto se aclara antes de la concentración de los virus influenza. Métodos comunes incluyen ultrafiltración, adsorción sobre sulfato de bario y elución, y centrifugación. Por ejemplo, el medio bruto procedente de cultivos infectados puede aclararse en primer lugar mediante centrifugación a, por ejemplo, 1000-2000 x g durante un tiempo suficiente para eliminar desechos celulares y otra materia particulada grande, por ejemplo, entre 10 y 30 minutos. Opcionalmente, luego se centrifuga el sobrenadante del medio aclarado para que sedimenten los virus influenza, por ejemplo, a 15.000 x g, durante aproximadamente 3-5 horas. Tras la suspensión del sedimento de virus en un tampón apropiado, tal como STE (Tris-HCl 0,01 M; NaCl 0,15 M; EDTA 0,0001 M) o solución salina tamponada con fosfato (PBS) a pH 7,4, se concentra el virus mediante centrifugación en gradiente de densidad sobre sacarosa (60%-12%) o tartrato de potasio (50%-10%). Son adecuados gradientes o bien continuos o bien escalonados, por ejemplo, un gradiente de sacarosa de entre el 12% y el 60% en cuatro etapas del 12%. Los gradientes se centrifugan a una velocidad y durante un tiempo suficientes para que los virus se concentren en una banda visible para su recuperación. Alternativamente, y para aplicaciones comerciales a mayor escala, el virus se somete a elutriación a partir de gradientes de densidad usando un rotor de centrifuga zonal en modo continuo. Detalles adicionales suficientes para guiar a un experto a través de la preparación de virus influenza procedentes de cultivo tisular se proporcionan, por ejemplo, en Furminger. *Vaccine Production*, en Nicholson *et al.* (eds.) *Textbook of Influenza*, págs. 324-332; Merten *et al.* (1996) *Production of influenza virus in cell cultures for vaccine preparation*, en Cohen & Shafferman (eds.) *Novel Strategies in Design and Production of Vaccines* págs. 141-151, y patente de Estados Unidos n.º 5.690.937. Si se desea, los virus recuperados pueden almacenarse a -80°C en presencia de sacarosa-fosfato-glutamato (SPG) como estabilizador.

Alternativamente, pueden emplearse sistemas de transcripción/traducción libres de células para producir polipéptidos que comprenden una secuencia o subsecuencia de aminoácidos o de, por ejemplo, las secuencias facilitadas en el presente documento tal como SEC ID N° 11-12, o codificarse mediante secuencias de polinucleótido, por ejemplo, SEC ID N° 1-2. Están comercialmente disponibles varios sistemas de transcripción y traducción *in vitro*. Una guía general para protocolos de transcripción y traducción *in vitro* se encuentra en Tymms (1995) *In vitro Transcription and Translation Protocols: Methods in Molecular Biology*, Volumen 37, Garland Publishing, NY.

Además, los polipéptidos, o subsecuencias de los mismos, por ejemplo, subsecuencias que comprenden péptidos antigénicos, pueden producirse manualmente o usando un sistema automatizado, mediante síntesis de péptidos directa usando técnicas en fase sólida (véase, Stewart *et al.* (1969) *Solid-Phase Peptide Synthesis*, WH Freeman Co, San Francisco; Merrifield J (1963) *J Am Chem Soc* 85:2149-2154). Los sistemas automatizados a modo de ejemplo incluyen el sintetizador de péptidos 431A de Applied Biosystems (Perkin Elmer, Foster City, CA). Si se desea, pueden sintetizarse químicamente subsecuencias por separado, y combinarse usando métodos químicos

para proporcionar polipéptidos de longitud completa.

#### Aminoácidos modificados

- 5 Los polipéptidos expresados descritos en el presente documento pueden contener uno o más aminoácidos modificados. La presencia de aminoácidos modificados puede ser ventajosa en, por ejemplo, (a) el aumento de la semivida en suero del polipéptido, (b) la reducción/aumento de la antigenicidad del polipéptido, (c) el aumento de la estabilidad en almacenamiento del polipéptido, etc. El(los) aminoácido(s) se modifican, por ejemplo, conjuntamente con la traducción o tras la traducción durante la producción recombinante (por ejemplo, glicosilación asociada al extremo N-terminal en motivos N-XS/T durante la expresión en células de mamífero) o se modifican mediante medios sintéticos (por ejemplo, a través de PEGilación).

Los ejemplos no limitantes de un aminoácido modificado incluyen un aminoácido glicosilado, un aminoácido sulfatado, un aminoácido prenilado (por ejemplo, farnesilado, geranilgeranilado), un aminoácido acetilado, un aminoácido acilado, un aminoácido pegilado, un aminoácido biotinilado, un aminoácido carboxilado, un aminoácido fosforilado, y similares, así como aminoácidos modificados mediante conjugación con, por ejemplo, restos lipídicos u otros agentes de derivatización orgánicos. Hay multitud de referencias bibliográficas adecuadas para guiar a un experto en la modificación de aminoácidos en la bibliografía. Los protocolos ejemplo se encuentran en Walker (1998) Protein Protocols on CD-ROM Human Press, Towata, NJ.

#### Proteínas de fusión

También se describen en el presente documento proteínas de fusión que comprenden fusiones de las secuencias descritas en el presente documento (por ejemplo, que codifican para polipéptidos de HA y/o NA tal como se facilita a modo de ejemplo mediante las SEC ID Nº 11-12) con, por ejemplo, inmunoglobulinas (o partes de las mismas), secuencias que codifican para, por ejemplo, GFP (proteína fluorescente verde), u otros marcadores similares, etc. Las proteínas de fusión se usan opcionalmente para, por ejemplo, aplicaciones similares (incluyendo, por ejemplo, aplicaciones terapéuticas, profilácticas, diagnósticas, experimentales, etc., tal como se describe en el presente documento) a las de las proteínas de no fusión descritas en el presente documento. Además de la fusión con secuencias de inmunoglobulina y secuencias de marcador, las proteínas descritas en el presente documento también se fusionan opcionalmente con, por ejemplo, secuencias que permiten la clasificación de las proteínas de fusión y/o el direccionamiento de las proteínas de fusión hacia regiones, tipos celulares específicos, etc.

#### Anticuerpos

35 Los polipéptidos descritos en el presente documento pueden usarse para producir anticuerpos específicos para los polipéptidos facilitados en el presente documento y/o polipéptidos codificados por los polinucleótidos descritos en el presente documento y variantes conservadoras de los mismos. Anticuerpos específicos para los polipéptidos mencionados anteriormente son útiles, por ejemplo, para fines de diagnóstico y terapéuticos, por ejemplo, relacionados con la actividad, distribución y expresión de polipéptidos diana.

Pueden generarse anticuerpos específicos para los polipéptidos descritos en el presente documento mediante métodos bien conocidos en la técnica. Tales anticuerpos pueden incluir, pero no se limitan a, policlonales, monoclonales, quiméricos, humanizados, de cadena sencilla, fragmentos Fab y fragmentos producidos por una biblioteca de expresión de Fab.

Los polipéptidos no requieren actividad biológica para la producción de anticuerpos (por ejemplo, no se requiere hemaglutinina o neuraminidasa funcionales de longitud completa). Sin embargo, el polipéptido u oligopéptido debe ser antigénico. Los péptidos usados para inducir anticuerpos específicos normalmente tienen una secuencia de aminoácidos de al menos aproximadamente 4 aminoácidos, y a menudo de al menos 5 ó 10 aminoácidos. Pueden fusionarse extensiones cortas de un polipéptido con otra proteína, tal como hemocianina de lapa californiana y un anticuerpo producido frente a la molécula química.

Los expertos en la técnica conocen numerosos métodos para producir anticuerpos policlonales y monoclonales y pueden adaptarse para producir anticuerpos específicos para los polipéptidos descritos en el presente documento y/o codificarse por la secuencia de polinucleótido descrita en el presente documento etc. Véase, por ejemplo, Coligan (1991) Current Protocols in Immunology, Wiley/Greene, NY; Paul (ed.) (1998) Fundamental Immunology, Cuarta Edición, Lippincott-Raven, Lippincott Williams & Wilkins; Harlow y Lane (1989) Antibodies: A Laboratory Manual Cold Spring Harbor Press, NY; Stites *et al.* (eds.) Basic and Clinical Immunology (4ª ed.) Lange Medical Publications, Los Altos, CA, y bibliografías citadas en ellos; Goding (1986) Monoclonal Antibodies: Principles and Practice (2ª ed.) Academic Press, Nueva York, NY; y Kohler y Milstein (1975) Nature 256: 495-497. Otras técnicas adecuadas para la preparación de anticuerpos incluyen la selección de bibliotecas de anticuerpos recombinantes en fagos o vectores similares. Véanse, Huse *et al.* (1989) Science 246: 1275-1281; y Ward, *et al.* (1989) Nature 341:544-546. Los antisueros y anticuerpos monoclonales y policlonales específicos se unirán habitualmente con una KD de, por ejemplo, al menos aproximadamente 0,1 µM, al menos aproximadamente 0,01 µM o mejor y normalmente al menos aproximadamente 0,001 µM o mejor.



Para ciertas aplicaciones terapéuticas, son deseables anticuerpos humanizados. Métodos detallados para la preparación de anticuerpos quiméricos (humanizados) pueden encontrarse en la patente estadounidense 5.482.856. Detalles adicionales sobre la humanización y otras técnicas de producción y modificación por ingeniería genética de anticuerpos pueden encontrarse en Borrebaeck (ed.) (1995) *Antibody Engineering*, 2ª Edición, Freeman and Company, NY (Borrebaeck); McCafferty *et al.* (1996) *Antibody Engineering, A Practical Approach* IRL en Oxford Press, Oxford, Inglaterra (McCafferty), y Paul (1995) *Antibody Engineering Protocols* Humana Press, Towata, NJ (Paul). Detalles adicionales en cuanto a procedimientos específicos pueden encontrarse, por ejemplo, en Ostberg *et al.* (1983), Hybridoma 2: 361-367, Ostberg, patente estadounidense n.º 4.634.664, y Engelman *et al.*, patente estadounidense n.º 4.634.666.

#### Definición de polipéptidos mediante inmunorreactividad

También se describen en el presente documento la generación de antisueros que se unen específicamente a los polipéptidos descritos en el presente documento así como los polipéptidos que se unen mediante tales antisueros.

Por ejemplo, se describen polipéptidos (por ejemplo, moléculas de HA y/o NA) que se unen específicamente a o que son específicamente inmunorreactivos con un anticuerpo o antisueros generados contra un inmunógeno que comprende una secuencia de aminoácidos seleccionada de una o más de las secuencias facilitadas en el presente documento (por ejemplo, SEC ID N° 11-12), etc. Para eliminar la reactividad cruzada con otros homólogos, el anticuerpo o antisuero se extraen con las moléculas de HA y/o NA encontradas en bases de datos públicas en el momento de presentar, por ejemplo, el(los) polipéptido(s) "control". Cuando las otras secuencias control corresponden a un ácido nucleico, se genera un polipéptido codificado por el ácido nucleico y se usa para fines de extracción de anticuerpo/antisueros.

En un formato típico, el inmunoensayo usa un antisuero policlonal que se produjo contra uno o más polipéptidos que comprenden una o más de las secuencias correspondientes a las secuencias en el presente documento (por ejemplo, SEC ID N° 11-12), etc. o una subsecuencia sustancial del mismo (es decir, al menos aproximadamente el 30% de la secuencia de longitud completa proporcionada). El conjunto de inmunógenos polipeptídicos potenciales derivados de las presentes secuencias se denomina colectivamente en lo sucesivo "polipéptidos inmunogénicos". El antisuero resultante se selecciona opcionalmente para que tenga baja reactividad cruzada contra homólogos de hemaglutinina y/o neuraminidasa control y se elimina cualquier reactividad cruzada de este tipo, por ejemplo, mediante inmunoabsorción, con uno o más de los homólogos de hemaglutinina y neuraminidasa control, antes del uso del antisuero policlonal en el inmunoensayo.

Con el fin de producir antisueros para su uso en un inmunoensayo, se produce y se purifica uno o más de los polipéptidos inmunogénicos tal como se describe en el presente documento. Por ejemplo, puede producirse una proteína recombinante en una célula recombinante. Se inmuniza una variedad endógama de ratones (usada en este ensayo porque los resultados son más reproducibles debido a la identidad genética virtual de los ratones) con la(s) proteína(s) inmunogénica(s) en combinación con un adyuvante convencional, tal como el adyuvante de Freund, y un protocolo de inmunización de ratones convencional (véase, por ejemplo, Harlow y Lane (1988) *Antibodies. A Laboratory Manual*, Cold Spring Harbor Publications, Nueva York, para una descripción convencional de generación de anticuerpos, formatos de inmunoensayos y condiciones que pueden usarse para determinar la inmunorreactividad específica). Pueden encontrarse bibliografía y comentarios adicionales de anticuerpos en el presente documento y pueden aplicarse en este caso para definir polipéptidos mediante inmunorreactividad. Alternativamente, uno o más polipéptido(s) sintético(s) o recombinante(s) derivados de las secuencias dadas a conocer en el presente documento se conjuga(n) con una proteína portadora y se usa(n) como inmunógeno.

Se recogen sueros policlonales y se titulan contra el polipéptido inmunogénico en un inmunoensayo, por ejemplo, un inmunoensayo en fase sólida con una o más de las proteínas inmunogénicas inmovilizadas sobre un soporte sólido. Se seleccionan antisueros policlonales con un título de  $10^6$  o mayor, se reúnen y se extraen con el(los) polipéptido(s) de hemaglutinina y/o neuraminidasa control para producir antisueros policlonales titulados, reunidos, extraídos.

Los antisueros policlonales titulados, reunidos, extraídos, se someten a prueba para determinar la reactividad cruzada contra el(los) homólogo(s) control en un inmunoensayo comparativo. En este ensayo comparativo, se determinan condiciones de unión discriminatorias para los antisueros policlonales, titulados, extraídos, lo que da como resultado al menos aproximadamente una razón de señal con respecto a ruido de 5-10 veces superior para la unión de los antisueros policlonales titulados a los polipéptidos inmunogénicos, en comparación con la unión a los homólogos control. Es decir, la rigurosidad de la reacción de unión se ajusta mediante la adición de competidores no específicos tales como albúmina o leche desnatada en polvo y/o mediante el ajuste de las condiciones salinas, la temperatura y/o similares. Estas condiciones de unión se usan en ensayos posteriores para determinar si un polipéptido de prueba (un polipéptido que se compara con los polipéptidos inmunogénicos y/o los polipéptidos control) se une específicamente mediante los antisueros policlonales extraídos reunidos. En particular, los polipéptidos de prueba que muestran al menos una razón de señal con respecto a ruido de 2-5x superior que los homólogos receptores control en condiciones de unión discriminatorias, y al menos aproximadamente un 1/2 de razón de señal con respecto a ruido en comparación con el(los) polipéptido(s) inmunogénico(s), comparten una similitud

estructural sustancial con el polipéptido inmunogénico en comparación con el receptor conocido, etc.

En otro ejemplo, se usan inmunoensayos en el formato de unión competitiva para la detección de un polipéptido de prueba. Por ejemplo, tal como se ha observado, se eliminan los anticuerpos de reactividad cruzada de la mezcla de antisueros reunidos mediante inmunoabsorción con los polipéptidos control. El(los) polipéptido(s) inmunogénico(s) se inmoviliza(n) entonces a un soporte sólido que se expone a los antisueros reunidos extraídos. Las proteínas de prueba se añaden al ensayo para competir por unirse a los antisueros reunidos extraídos. La capacidad de la(s) proteína(s) de prueba para competir por unirse a los antisueros reunidos extraídos en comparación con la(s) proteínas(s) inmovilizada(s) se compara con la capacidad del(de los) polipéptido(s) inmunogénico(s) añadido(s) al ensayo para competir por unirse (los polipéptidos inmunogénicos compiten eficazmente con los polipéptidos inmunogénicos inmovilizados por unirse a los antisueros reunidos). Se calcula el porcentaje de reactividad cruzada para las proteínas de prueba, usando cálculos convencionales.

En un ensayo paralelo, se determina opcionalmente la capacidad de la(s) proteína(s) control para competir por unirse a los antisueros reunidos extraídos en comparación con la capacidad del(de los) polipéptido(s) inmunogénico(s) para competir por unirse a los antisueros. De nuevo, se calcula el porcentaje de reactividad cruzada para el(los) polipéptido(s) control, usando cálculos convencionales. Cuando el porcentaje de reactividad cruzada es al menos 5-10x tan alto para los polipéptidos de prueba en comparación con el(los) polipéptido(s) control y/o cuando la unión de los polipéptidos de prueba está aproximadamente en el intervalo de la unión de los polipéptidos inmunogénicos, se dice que los polipéptidos de prueba se unen específicamente a los antisueros reunidos.

En general, los antisueros inmunoabsorbidos y reunidos pueden usarse en un inmunoensayo de unión competitiva tal como se describe en el presente documento para comparar cualquier polipéptido de prueba con el(los) polipéptido(s) inmunogénico(s) y/o control. Con el fin de realizar esta comparación, se someten a ensayo los polipéptidos inmunogénicos, de prueba y control en un amplio intervalo de concentraciones y se determina la cantidad de cada polipéptido requerida para inhibir el 50% de la unión de los antisueros reunidos a, por ejemplo, una proteína inmovilizada inmunogénica, de prueba o control, usando técnicas convencionales. Si la cantidad del polipéptido de prueba requerida en el ensayo competitivo es menos de dos veces la cantidad del polipéptido inmunogénico que se requiere, entonces se dice que el polipéptido de prueba se une específicamente a un anticuerpo generado frente a la proteína inmunogénica, siempre que la cantidad sea al menos aproximadamente 5-

10x tan alta como para el polipéptido control.

Como una determinación adicional de la especificidad, el antisuero reunido se inmunoabsorbe opcionalmente de manera completa con el(los) polipéptido(s) inmunogénico(s) (en lugar de con el(los) polipéptido(s) control) hasta que se detecte poca o ninguna unión de los antisueros reunidos extraídos con polipéptido inmunogénico resultantes al(a los) polipéptido(s) inmunogénico(s) usado(s) en la inmunoabsorción. Este antisuero completamente inmunoabsorbido se somete entonces a prueba para determinar la reactividad con el polipéptido de prueba. Si se observa poca o ninguna reactividad (es decir, no más de 2x la razón de señal con respecto a ruido observada para la unión de los antisueros completamente inmunoabsorbidos al polipéptido inmunogénico), entonces el polipéptido de prueba se une específicamente mediante los antisueros provocados por la proteína inmunogénica.

#### VARIANTES DE SECUENCIA DE ÁCIDO NUCLEICO Y POLIPÉPTIDO

##### Variaciones silenciosas

Debido a la degeneración del código genético, se produce opcionalmente cualquiera de una variedad de secuencias de ácido nucleico que codifican para los polipéptidos descritos en el presente documento, pudiendo portar algunas de las ellas niveles inferiores de identidad de secuencia con las secuencias de ácido nucleico y polipéptido de HA y NA en el presente documento. A continuación se facilita una tabla de codones típicos que especifica el código genético, encontrada en muchos textos de biología y bioquímica.

Tabla 1

| Tabla de codones |     |   |       |     |     |     |
|------------------|-----|---|-------|-----|-----|-----|
| Aminoácidos      |     |   | Codón |     |     |     |
| Alanina          | Ala | A | GCA   | GCC | GCG | GCU |
| Cisteína         | Cys | C | UGC   | UGU |     |     |
| Ácido aspártico  | Asp |   | GAC   | GAU |     |     |
| Ácido glutámico  | Glu |   | GAA   | GAG |     |     |
| Fenilalanina     | Phe | F | UUC   | UUU |     |     |

|            |     |   |     |     |     |     |     |     |  |
|------------|-----|---|-----|-----|-----|-----|-----|-----|--|
| Glicina    | Gly | G | GGA | GGC | GGG | GGU |     |     |  |
| Histidina  | His | H | CAC | CAU |     |     |     |     |  |
| Isoleucina | Ile | I | AUA | AUC | AUU |     |     |     |  |
| Lisina     | Lys | K | AAA | AAG |     |     |     |     |  |
| Leucina    | Leu | L | UUA | UUG | CUA | CUC | CUG | CUU |  |
| Metionina  | Met | M | AUG |     |     |     |     |     |  |
| Asparagina | Asn | N | AAC | AAU |     |     |     |     |  |
| Prolina    | Pro | P | CCA | CCC | CCG | CCU |     |     |  |
| Glutamina  | Gln | Q | CAA | CAG |     |     |     |     |  |
| Arginina   | Arg | R | AGA | AGG | CGA | CGC | CGG | CGU |  |
| Serina     | Ser | S | AGC | AGU | UCA | UCC | UCG | UCU |  |
| Treonina   | Thr | T | ACA | ACC | ACG | ACU |     |     |  |
| Valina     | Val | V | GUA | GUC | GUG | GUU |     |     |  |
| Triptófano | Trp | W | UGG |     |     |     |     |     |  |
| Tirosina   | Tyr | Y | UAC | UAU |     |     |     |     |  |

La tabla de codones muestra que muchos aminoácidos están codificados por más de un codón. Por ejemplo, los codones AGA, AGG, CGA, CGC, CGG y CGU codifican todos para el aminoácido arginina. Por tanto, en cada posición en los ácidos nucleicos descritos en el presente documento en los que una arginina está especificada por un codón, el codón puede alterarse a cualquiera de los codones correspondientes descritos anteriormente sin alterar el polipéptido codificado. Se entiende que U en una secuencia de ARN se corresponde con T en una secuencia de ADN.

Tales “variaciones silenciosas” son una especie de “variaciones modificadas de manera conservadora” comentadas a continuación. Un experto reconocerá que cada codón en un ácido nucleico (excepto ATG, que normalmente es el único codón para metionina, y TTG, que normalmente es el único codón para triptófano) puede modificarse mediante técnicas convencionales de manera que codifique para un polipéptido funcionalmente idéntico. Por consiguiente, cada variación silenciosa de un ácido nucleico que codifica para un polipéptido está implícita en cualquier secuencia descrita. La descripción, por tanto, proporciona explícitamente todas y cada una de las posibles variaciones de una secuencia de ácido nucleico que codifica para un polipéptido descrito en el presente documento que podrían realizarse seleccionando combinaciones basadas en posibles elecciones de codones. Estas combinaciones se realizan según el código genético de tripletes convencional (por ejemplo, tal como se expone en la Tabla 1, o tal como está comúnmente disponible en la técnica) tal como se aplica a la secuencia de ácido nucleico que codifica para un polipéptido de hemaglutinina o neuraminidasa descrito en el presente documento. Todas estas variaciones de cada ácido nucleico en el presente documento se proporcionan específicamente y se describen teniendo en cuenta la secuencia en combinación con el código genético. Un experto en la técnica es completamente capaz de realizar estas sustituciones silenciosas usando los métodos en el presente documento.

#### Variaciones conservadoras

Debido a la degeneración del código genético, las “sustituciones silenciosas” (es decir, sustituciones en una secuencia de ácido nucleico que no dan como resultado una alteración en un polipéptido codificado) son una característica implícita de cada secuencia de ácido nucleico descrita en el presente documento que codifica para un aminoácido.

Las “variaciones conservadoras” de una secuencia de ácido nucleico particular se refieren a aquellos ácidos nucleicos que codifican para secuencias de aminoácidos idénticas. Finalmente, la adición de secuencias que no alteran la actividad codificada de una molécula de ácido nucleico, tal como la adición de una secuencia no funcional, es una variación conservadora del ácido nucleico básico.

#### Subsecuencias únicas de polipéptido y polinucleótido

También se describe en el presente documento un ácido nucleico que comprende una subsecuencia única en un ácido nucleico seleccionada de la secuencia de moléculas de HA y/o NA dada a conocer en el presente documento, por ejemplo, SEC ID N° 1-2. La subsecuencia única es única en comparación con un ácido nucleico correspondiente a ácidos nucleicos tales como, por ejemplo, los encontrados en GenBank u otras bases de datos públicas similares en el momento de la presentación. La alineación puede realizarse usando, por ejemplo, la herramienta BLAST

configurada con parámetros por defecto. Cualquier subsecuencia única es útil, por ejemplo, como sonda para identificar los ácidos nucleicos descritos en el presente documento. Véase anteriormente.

- De manera similar, se describe en el presente documento un polipéptido que comprende una subsecuencia única en un polipéptido seleccionada de la secuencia de moléculas de HA y/o NA dada a conocer en el presente documento, por ejemplo, SEC ID N° 11-12. En este caso, la subsecuencia única es única en comparación con un polipéptido correspondiente a, por ejemplo, el aminoácido correspondiente a las secuencias de polinucleótido encontradas en, por ejemplo, GenBank u otras bases de datos públicas similares en el momento de la presentación.
- Además se describen ácidos nucleicos diana que hibridan en condiciones rigurosas con un oligonucleótido codificante único que codifica para una subsecuencia única en un polipéptido seleccionada de las secuencias de moléculas de HA y/o NA descritas en el presente documento, siendo la subsecuencia única en comparación con un polipéptido correspondiente a cualquiera de los polipéptidos de control (secuencias de, por ejemplo, los ácidos nucleicos correspondientes a los encontrados en, por ejemplo, GenBank u otras bases de datos públicas similares en el momento de la presentación). Las secuencias únicas se determinan tal como se indicó anteriormente.

#### Comparación, identidad y homología de secuencias

- Los términos “idéntico” o “identidad” en porcentaje, en el contexto de dos o más secuencias de ácido nucleico o polipéptido, se refieren a dos o más secuencias o subsecuencias que son iguales o tienen un porcentaje especificado de residuos de aminoácidos o nucleótidos que son iguales, cuando se comparan y alinean para obtener una correspondencia máxima, tal como se mide usando uno de los algoritmos de comparación de secuencias descritos más adelante (u otros algoritmos disponibles para los expertos) o mediante inspección visual.
- La expresión “sustancialmente idéntico,” en el contexto de dos ácidos nucleicos o polipéptidos (por ejemplo, ADN que codifican para una molécula de HA o NA, o la secuencia de aminoácidos de una molécula de HA o NA) se refiere a dos o más secuencias o subsecuencias que tienen una identidad de residuos de nucleótido o aminoácido de al menos aproximadamente el 90%, preferiblemente el 91%, lo más preferiblemente el 92%, 93%, 94%, 95%, 96%, 97%, 98%, 98,5%, 99%, 99,1%, 99,2%, 99,3%, 99,4%, 99,5%, 99,6%, 99,7%, 99,8%, 99,9% o más, cuando se comparan y alinean para obtener una correspondencia máxima, tal como se mide usando un algoritmo de comparación de secuencias o mediante inspección visual. Normalmente se considera que tales secuencias “sustancialmente idénticas” son homólogas, sin hacer referencia a los ancestros reales. Preferiblemente, existe “identidad sustancial” a lo largo de una región de las secuencias de aminoácidos que tiene al menos aproximadamente 200 residuos de longitud, más preferiblemente a lo largo de una región de al menos aproximadamente 250 residuos, y lo más preferiblemente las secuencias son sustancialmente idénticas a lo largo de al menos aproximadamente 300 residuos, 350 residuos, 400 residuos, 425 residuos, 450 residuos, 475 residuos, 480 residuos, 490 residuos, 495 residuos, 499 residuos, 500 residuos, 502 residuos, 559 residuos, 565 residuos, o a lo largo de la longitud completa de las dos secuencias que van a compararse.
- Para la comparación de secuencias y la determinación de la homología, normalmente una secuencia actúa como secuencia de referencia con la que se comparan las secuencias de prueba. Cuando se usa un algoritmo de comparación de secuencias, se introducen las secuencias de prueba y de referencia en un ordenador, se designan coordenadas de subsecuencia si es necesario, y se diseñan los parámetros del programa de algoritmo de secuencias. El algoritmo de comparación de secuencias calcula entonces la identidad de secuencias en porcentaje para la(s) secuencia(s) de prueba con respecto a la secuencia de referencia, basándose en los parámetros designados del programa.

- La alineación óptima de secuencias para la comparación puede realizarse, por ejemplo, mediante el algoritmo de homología local de Smith & Waterman, Adv Appl Math 2:482 (1981), mediante el algoritmo de alineación por homología de Needleman & Wunsch, J Mol Biol 48:443 (1970), mediante el método de búsqueda de similitud de Pearson & Lipman, Proc Natl Acad Sci USA 85:2444 (1988), mediante implementaciones informatizadas de algoritmos tales como GAP, BESTFIT, FASTA y TFASTA en el paquete de software Wisconsin Genetics, Genetics Computer Group, 575 Science Dr., Madison, WI, o mediante inspección visual (véase generalmente, Ausubel *et al.*, citado anteriormente).

- Un ejemplo de un algoritmo que es adecuado para determinar la similitud de secuencia y la identidad de secuencia en porcentaje es el algoritmo BLAST, que se describe en Altschul *et al.*, J Mol Biol 215:403-410 (1990). El software para realizar los análisis de BLAST está públicamente disponible a través del National Center for Biotechnology Information ([www.ncbi.nlm.nih.gov/](http://www.ncbi.nlm.nih.gov/)). Este algoritmo implica identificar en primer lugar pares de secuencias de alta puntuación (HSP) identificando palabras cortas de longitud W en la secuencia problema, que o bien corresponden o bien satisfacen alguna puntuación T umbral de valor positivo cuando se alinean con una palabra de la misma longitud en una secuencia de la base de datos. T se refiere al umbral de puntuación de palabras vecinas (véase, Altschul *et al.*, citado anteriormente). Estos aciertos de palabras vecinas iniciales actúan como semillas para iniciar búsquedas para encontrar HSP más largos que las contienen. Los aciertos de palabras se extienden entonces en ambas direcciones a lo largo de cada secuencia tanto como pueda aumentarse la puntuación de alineación acumulativa. Las puntuaciones acumulativas se calculan usando, para secuencias de nucleótidos, los parámetros M

(puntuación de recompensa para un par de residuos coincidentes; siempre  $> 0$ ) y N (puntuación de penalización para residuos no coincidentes; siempre  $< 0$ ). Para secuencias de aminoácidos, se usa una matriz de puntuación para calcular la puntuación acumulativa. La extensión de los aciertos de palabras en cada dirección se detiene cuando: la puntuación de alineación acumulativa disminuye por debajo de la cantidad X desde su valor máximo conseguido; la puntuación acumulativa llega hasta cero o menos, debido a la acumulación de una o más alineaciones de residuos de puntuación negativa; o se alcanza el final de cualquier secuencia. Los parámetros W, T y X del algoritmo BLAST determinan la sensibilidad y velocidad de la alineación. El programa BLASTN (para secuencias de nucleótidos) usa como parámetros por defecto una longitud de palabra (W) de 11, una expectativa (E) de 10, un punto de corte de 100, M=5, N=4, y una comparación de ambas cadenas. Para secuencias de aminoácidos, el programa BLASTP usa como parámetros por defecto una longitud de palabra (W) de 3, una expectativa (E) de 10, y la matriz de puntuación BLOSUM62 (véase, Henikoff & Henikoff (1989) Proc Natl Acad Sci USA 89:10915).

Además de calcular la identidad de secuencia en porcentaje, el algoritmo BLAST también realiza un análisis estadístico de la similitud entre dos secuencias (véase, por ejemplo, Karlin & Altschul, Proc Natl Acad Sci USA 90:5873-5877 (1993)). Una medida de la similitud proporcionada por el algoritmo BLAST es la probabilidad de suma más pequeña (P(N)), que proporciona una indicación de la probabilidad por la que una coincidencia entre dos secuencias de nucleótidos o aminoácidos se produciría por casualidad. Por ejemplo, un ácido nucleico se considera similar a una secuencia de referencia si la probabilidad de suma más pequeña en una comparación del ácido nucleico de prueba con el ácido nucleico de referencia es inferior a aproximadamente 0,1, más preferiblemente inferior a aproximadamente 0,01, y lo más preferiblemente inferior a aproximadamente 0,001.

Otro ejemplo de un algoritmo de alineación de secuencias útil es PILEUP. PILEUP crea una alineación de secuencias múltiple a partir de un grupo de secuencias relacionadas usando alineaciones progresivas, por parejas. También puede representarse gráficamente un árbol que muestra las relaciones de agrupación usadas para crear la alineación. PILEUP usa una simplificación del método de alineación progresiva de Feng & Doolittle (1987) J. Mol. Evol. 35:351-360. El método usado es similar al método descrito por Higgins & Sharp (1989) CABIOS5:151-153. El programa puede alinear, por ejemplo, hasta 300 secuencias de una longitud máxima de 5.000 letras. El procedimiento de alineación múltiple comienza con la alineación por parejas de las dos secuencias más similares, produciendo una agrupación de dos secuencias alineadas. Esta agrupación puede alinearse entonces con la siguiente secuencia más relacionada o agrupación de secuencias alineadas. Dos agrupaciones de secuencias pueden alinearse mediante una simple extensión de la alineación por parejas de dos secuencias individuales. La alineación final se logra mediante una serie de alineaciones progresivas, por parejas. El programa también puede usarse para representar gráficamente un dendograma o representación en árbol de relaciones de agrupación. El programa se ejecuta designando secuencias específicas y sus coordenadas de aminoácidos o nucleótidos para regiones de comparación de secuencias.

Un ejemplo adicional de un algoritmo que es adecuado para múltiples alineaciones de secuencias de aminoácidos o ADN es el programa CLUSTALW (Thompson, J. D. *et al.* (1994) Nucl. Acids. Res. 22: 4673-4680). CLUSTALW realiza múltiples comparaciones por parejas entre grupos de secuencias y las reúne en una alineación múltiple basada en la homología. Las penalizaciones por apertura de hueco y extensión de hueco pueden ser, por ejemplo, de 10 y 0,05, respectivamente. Para alineaciones de aminoácidos, puede usarse el algoritmo BLOSUM como matriz de pesos de proteínas. Véase, por ejemplo, Henikoff and Henikoff (1992) Proc. Natl. Acad. Sci. USA 89: 10915-10919.

## SISTEMAS DIGITALES

Se describen adicionalmente sistemas digitales, por ejemplo, ordenadores, medios legibles por ordenador y sistemas integrados que comprenden cadenas de caracteres correspondientes a la información de secuencia en el presente documento para los ácidos nucleicos y polipéptidos aislados o recombinantes en el presente documento, incluyendo, por ejemplo, las secuencias mostradas en el presente documento, y las diversas sustituciones silenciosas y sustituciones conservadoras de las mismas. Los sistemas integrados pueden incluir además, por ejemplo, equipo de síntesis de genes para obtener genes correspondientes a las cadenas de caracteres.

Pueden usarse diversos métodos conocidos en la técnica para detectar homología o similitud entre diferentes cadenas de caracteres (véase anteriormente), o pueden usarse para realizar otras funciones deseables tales como para controlar archivos de salida, proporcionar la base para realizar presentaciones de información incluyendo las secuencias y similares. Los ejemplos incluyen BLAST, comentado anteriormente. Los sistemas informáticos pueden incluir tales programas, por ejemplo, conjuntamente con uno o más archivo(s) de datos o base(s) de datos que comprende(n) una secuencia tal como se indica en el presente documento.

Por tanto, pueden detectarse y reconocerse diferentes tipos de homología y similitud de diversa rigurosidad y longitud entre diversas secuencias o fragmentos de HA o NA, etc. en los sistemas integrados en el presente documento. Por ejemplo, se han diseñado muchos métodos de determinación de la homología para el análisis comparativo de secuencias de biopolímeros, para la "corrección ortográfica" en el procesamiento de textos y para la recuperación de datos de diversas bases de datos. Con comprensión de las interacciones complementarias por parejas de la doble hélice entre 4 nucleobases principales en los polinucleótidos naturales, también pueden usarse

modelos que simulan la hibridación de cadenas de polinucleótidos homólogos complementarios como base de la alineación de secuencias u otras operaciones realizadas normalmente en las cadenas de caracteres correspondientes a las secuencias en el presente documento (por ejemplo, manipulaciones del procesamiento de palabras, construcción de figuras que comprenden cadenas de caracteres de secuencia o subsecuencia, tablas de salida, etc.).

Por tanto, pueden adaptarse aplicaciones de sobremesa convencionales tales como software de procesamiento de textos (por ejemplo, Microsoft Word™ o Corel WordPerfect™) y software de bases de datos (por ejemplo, software de hojas de cálculo tal como Microsoft Excel™, Corel Quattro Pro™, o programas de bases de datos tales como Microsoft Access™, Paradox™, GeneWorks™, o MacVector™ u otros programas similares) introduciendo una cadena de caracteres correspondiente a uno o más polinucleótidos y polipéptidos descritos en el presente documento (o bien ácidos nucleicos o bien proteínas, o ambos). Por ejemplo, un sistema puede incluir el software anterior que tiene la información de cadena de caracteres apropiada, por ejemplo, usado conjuntamente con una interfaz de usuario (por ejemplo, una GUI en un sistema operativo convencional tal como un sistema de Windows, Macintosh o LINUX) para manipular cadenas de caracteres correspondientes a las secuencias en el presente documento. Tal como se ha indicado, también pueden incorporarse programas de alineación especializados tales como BLAST en los sistemas descritos en el presente documento para la alineación de ácidos nucleicos o proteínas (o cadenas de caracteres correspondientes).

Los sistemas descritos en el presente documento incluyen normalmente un ordenador digital con conjuntos de datos introducidos en el sistema de software que comprenden cualquiera de las secuencias en el presente documento. El ordenador puede ser, por ejemplo, un PC (máquina basada en DOS™, OS2™, WINDOWS™, WINDOWSNT™, WINDOWS95™, WINDOWS2000™, WINDOWS98™, LINUX compatible con un procesador Intel x86 o Pentium, una máquina basada en MACINTOSH™, Power PC o UNIX (por ejemplo, la estación de trabajo SUN™) u otro ordenador disponible comercialmente que conozca el experto. Se dispone del software para alinear o manipular de otro modo secuencias, o puede construirse fácilmente por un experto usando un lenguaje de programación convencional tal como Visualbasic, PERL, Fortran, Basic, Java o similares.

Cualquier controlador u ordenador incluye opcionalmente un monitor que a menudo es una pantalla de tubo de rayos catódicos ("CRT"), una pantalla de panel plano (por ejemplo, pantalla de cristal líquido de matriz activa, pantalla de cristal líquido), u otras. El conjunto de circuitos del ordenador a menudo se coloca en una caja que incluye numerosos chips de circuito integrado, tales como un microprocesador, memoria, circuitos de interfaz y otros. La caja también incluye opcionalmente una unidad de disco duro, una unidad de disco flexible, una unidad extraíble de alta capacidad tal como un CD-ROM grabable, y otros elementos periféricos comunes. Dispositivos de entrada tales como un teclado o un ratón proporcionan opcionalmente la introducción por parte de un usuario y que el usuario compare la selección de secuencias o las manipule de otro modo en el sistema informático relevante.

El ordenador incluye normalmente software apropiado para recibir instrucciones del usuario, o bien en forma de entrada de usuario en un campo de parámetros fijados, por ejemplo, en una GUI, o bien en forma de instrucciones programadas previamente, por ejemplo, programadas previamente para una variedad de operaciones específicas diferentes. El software convierte entonces estas instrucciones en lenguaje apropiado para instruir el funcionamiento, por ejemplo, de mecanismos o controladores de transporte apropiados para llevar a cabo la operación deseada. El software también puede incluir elementos de salida para controlar la síntesis de ácidos nucleicos (por ejemplo, basándose en una secuencia o una alineación de secuencias en el presente documento); comparaciones de muestras para la expresión génica diferencial, u otras operaciones.

#### KITS Y REACTIVOS

También se describe un kit. Por ejemplo, un kit contiene uno o más ácido(s) nucleico(s), polipéptido(s), anticuerpo(s) o línea(s) celular(es) descritos en el presente documento (por ejemplo, que comprenden, o con, una molécula de HA y/o NA descrita en el presente documento). El kit puede contener un ácido nucleico o polipéptido de diagnóstico, por ejemplo, un anticuerpo, conjunto de sonda, por ejemplo, como un microalineamiento de ADNc empaquetado en un envase adecuado, u otro ácido nucleico tal como uno o más vector(es) de expresión. El kit también puede comprender adicionalmente uno o más reactivos adicionales, por ejemplo, sustratos, marcadores, cebadores, para marcar productos de expresión, tubos y/u otros accesorios, reactivos para recoger muestras, tampones, cámaras de hibridación, cubreobjetos, etc. El kit comprende opcionalmente además un conjunto de instrucciones o manual de usuario que detalla los métodos preferidos de uso de los componentes del kit para el descubrimiento o la aplicación de conjuntos de diagnóstico, etc.

Cuando se usa según las instrucciones, el kit puede usarse, por ejemplo, para evaluar un estado o patología, para evaluar los efectos de un agente farmacéutico u otra intervención del tratamiento en la progresión de un estado o patología en una célula u organismo, o para su uso como vacuna, etc.

Se describen adicionalmente kits de sistema que incorporan los métodos, la composición, los sistemas y el aparato en el presente documento. Los kits de sistema comprenden opcionalmente uno o más de los siguientes: (1) un aparato, sistema, componente de sistema o componente de aparato; (2) instrucciones para poner en práctica los

métodos descritos en el presente documento, y/o para hacer funcionar el aparato o componentes del aparato en el presente documento y/o para usar las composiciones en el presente documento. También se describe el uso de cualquier aparato, componente del aparato, composición o kit en el presente documento, para la puesta en práctica de cualquier método o ensayo en el presente documento, y/o para el uso de cualquier aparato o kit para poner en práctica cualquier ensayo o método en el presente documento.

Adicionalmente, los kits pueden incluir uno o más sistema(s) de traducción tal como se indicó anteriormente (por ejemplo, una célula) con material de empaquetamiento apropiado, envases para contener los componentes del kit, materiales de instrucciones para poner en práctica los métodos en el presente documento y/o similares. De manera similar, pueden proporcionarse productos de los sistemas de traducción (por ejemplo, proteínas tales como moléculas de HA y/o NA) en forma de kit, por ejemplo, con envases para contener los componentes del kit, materiales de instrucciones para poner en práctica los métodos en el presente documento y/o similares.

Para facilitar el uso de los métodos y las composiciones de la invención, pueden empaquetarse en forma de kit cualquiera de los componentes y/o composiciones de vacuna, por ejemplo, virus reagrupado en líquido alantoico, etc., y componentes adicionales, tales como, tampón, células, medio de cultivo, útiles para el empaquetamiento y la infección de los virus influenza para fines de vacuna experimental o terapéutica. Normalmente, el kit contiene, además de los componentes anteriores, materiales adicionales que pueden incluir, por ejemplo, instrucciones para realizar los métodos de la invención, material de empaquetamiento y un envase.

## Ejemplos

### Construcción y análisis de vacunas y virus H5N1 *ca*

Se usaron diversas secuencias en el presente documento que comprenden las secuencias de HA/NA H5N1 para crear vacunas y virus influenza. Se alteraron las secuencias de HA en tales vacunas con respecto al tipo natural mediante la eliminación del sitio de escisión polibásico dentro de la HA. Se reagruparon las secuencias de HA/NA (en un reagrupamiento 6:2) con A/AA/6/60 (un virus *att*, *ca*, véase anteriormente).

Se usaron tres cepas de influenza H5N1 en este ejemplo: A/VN/1203/2004, A/HK/491/97 y A/HK/213/2003. Tales cepas también se denominan dentro de este ejemplo las cepas del 97, 03 y 04 basándose en las designaciones de sus años. La similitud en porcentaje de los genes de HA de estas tres cepas es del 95-96%. La figura 1 ilustra la modificación del sitio de escisión polibásico de una secuencia de HA a modo de ejemplo, las secuencias de HA del 04, usadas para construir los virus/vacunas. Tal como se estableció anteriormente, diversas realizaciones de la invención comprenden secuencias que tienen diferentes regiones del sitio de escisión polibásico eliminadas. Véase anteriormente.

Tal como se ha establecido, se usaron las secuencias de H5N1 modificadas (es decir, los genes del 97, 03 y 04 modificados) para construir virus con reagrupamiento 6:2 con A/AA/6/60. Se apreciará y se señala en otra parte en el presente documento que también podrían haberse usado otras estructuras principales deseables (por ejemplo, PR8, etc.).

En los reagrupamientos 6:2 de este ejemplo, las secuencias de genes de HA y NA se derivaron del virus original de tipo natural y los genes restantes se caracterizaron mediante análisis de secuencia derivado del virus original A/AA/6/60 *ca*. Los virus reagrupados se replicaron hasta 8,0-8,5 log<sub>10</sub>DICT<sub>50</sub> en huevos. Sin embargo, se apreciará que un virus tal como se reivindica en el que el log<sub>10</sub> DICT<sub>50</sub> está comprendido entre aproximadamente 7,0 y aproximadamente 9,0, entre aproximadamente 7,5 y 8,5, o entre aproximadamente 8,0 y 8,5 también forma parte de la invención. Se limitó la capacidad de escisión de la HA modificada en los virus construidos por proteasas endógenas *in vitro* y los virus eran dependientes de tripsina (por ejemplo, desde aproximadamente 0,1 ug/ml hasta aproximadamente 1,0 ug/ml) para el crecimiento. Los virus construidos eran sensibles a la temperatura *in vitro*.

Los virus reagrupados H5N1 *ca* (que tienen los genes de HA del 97, 03 o 04 modificados) no fueron altamente patógenos para pollos. Por ejemplo, cuando se inocularon pollos Plymouth Rock blancos SPF de 4 semanas de edad por vía intravenosa con una dilución 1:10 de virus de reserva (10<sup>8-8,75</sup> DICT<sub>50</sub>/ml) y se observaron durante 10 días, se observó que 8 de 8 pollos murieron en el plazo de 1-2 días cuando se usaron H5N1 97, 03 y 04 de tipo natural, mientras que murieron 0 de 8 pollos cuando se usaron virus reagrupados H5N1 *ca*. Tal como puede observarse en la figura 2, los virus reagrupados H5N1 *ca* administrados por vía intranasal no se replicaron en pollos.

Los virus reagrupados H5N1/AA *ca* tampoco fueron letales para los ratones. Véase la figura 3, que también muestra la DICT<sub>50</sub> para las cepas de tipo natural H5N1. La figura 4 muestra que los virus reagrupados H5N1 *ca* 1997 y 2004 estaban restringidos en la replicación en ratones. La figura 5 muestra que los virus reagrupados H5N1 *ca* estaban restringidos en la replicación en pulmones de ratones.

En la figura 6 se muestra una comparación de los títulos de anticuerpos HAI séricos producidos en ratones tras una única dosis intranasal de vacuna (2003 *ca* en comparación con 2003 de tipo natural).

La figura 7 muestra mediciones similares, pero usando títulos de anticuerpos neutralizantes séricos.

La figura 8 muestra que los virus reagrupados H5N1 *ca* protegen a ratones de la exposición letal con 50, 500 ó 5.000 DL<sub>50</sub> del virus H5N1 de tipo natural. La figura 9 muestra la eficacia de protección frente a la replicación pulmonar de virus de exposición H5N1 homólogos y heterólogos en ratones. Tal como puede observarse, los virus reagrupados *ca* se replicaron peor que los virus de tipo natural. La figura 10 muestra datos relacionados usando aparatos respiratorios superiores de ratones. Los expertos en la técnica estarán familiarizados con exposiciones de homólogos y heterólogos (por ejemplo, someter a prueba si la vacuna del 2003 protege frente a una exposición del 2003 de tipo natural (homólogo) o si una vacuna del 2003 protege frente a una exposición del 1997 de tipo natural (heterólogo), etc.).

La figura 11 muestra la eficacia de protección conferida por la vacuna H5N1 *ca* 2004 frente a la exposición a alta dosis (10<sup>5</sup>DICT<sub>50</sub>) con virus H5N1 de tipo natural homólogos o heterólogos en ratones. La figura 12 muestra la eficacia de protección conferida por las vacunas H5N1 *ca* 1997 y 2003 frente a la exposición a alta dosis (10<sup>5</sup>DICT<sub>50</sub>) con virus de tipo natural H5N1 homólogos o heterólogos en ratones. La figura 13 muestra la eficacia de protección conferida por la vacuna H5N1 *ca* 2004 frente a dosis bajas o altas de exposición a virus de tipo natural H5N1 homólogos en ratones. Las figuras 11-13 demuestran que las vacunas sometidas a prueba podrían proteger frente a otros virus relacionados.

El ejemplo actual demuestra varios puntos relativos a vacunas/virus reagrupados H5N1 *ca* a modo de ejemplo de la invención. Se demostró que los virus reagrupados *ca* 97, 03 y 04 tienen fenotipo *ts in vitro*, pérdida de patogenicidad en pollos y atenuación en ratones. Se espera que la atenuación también esté presente en hurones. También se demostró la eficacia de protección y protección cruzada frente a la exposición letal y la propagación sistémica con virus de tipo natural en ratones. También se espera eficacia de protección y protecciones cruzadas frente a la replicación de virus de exposición de tipo natural en el aparato respiratorio de ratones.

Se contempla usar estos virus/vacunas (y similares) para determinar si la inmunogenicidad y la eficacia mejoran tras 2 dosis de vacuna; para evaluar la inmunogenicidad en primates no humanos; para evaluar la atenuación y la eficacia de la vacuna en hurones; para determinar la contribución de la inmunidad humoral y celular a la eficacia observada de las vacunas producidas en ratones; para determinar qué residuos de la HA 2003 contribuyen a la inmunogenicidad potenciada e introducirlos en las HA 1997 y 2004; y para determinar los efectos de delecionar el sitio de escisión de aminoácidos multibásico y de la constelación de genes.

Continúa en página siguiente



## SECUENCIAS

A/Vietnam/1203/04

Secuencia de nucleótidos de A/Vietnam/1203/04 H5 (SEC ID N° 1) Longitud de la molécula completa: 1767 nt

```

1 agcaaaagca ggggttcaat ctgtcaaaat ggagaaaata gtgcttcttt
51 ttgcaatagt cagtcttggt aaaagtgatc agatttgcat tggttaccat
101 gcaaacaact cgacagagca ggttgacaca ataatggaaa agaacgttac
151 tgttacacat gcccaagaca tactggaaaa gaaacacaac ggggaagctct
201 gcgatctaga tggagtgaag cctctaattt tgagagattg tagcgtagct
251 ggatggctcc tcggaaaccc aatgtgtgac gaattcatca atgtgccgga
301 atggtcttac atagtggaga aggccaatcc agtcaatgac ctctgttacc
351 caggggattt caatgactat gaagaattga aacacctatt gagcagaata
401 aaccattttg agaaaattca gatcatcccc aaaagttctt ggtccagtca
451 tgaagcctca ttaggggtga gctcagcatg tccataccag ggaaagtcct
501 cctttttcag aaatgtggta tggcttatca aaaagaacag tacataccca
551 acaataaaga ggagctacaa taataccaac caagaagatc ttttggtact
601 gtgggggatt caccatccta atgatgcggc agagcagaca aagctctatc
651 aaaacccaac cacctatatt tccgttgga catcaacact aaaccagaga
701 ttggtaccaa gaatagctac tagatccaaa gttaaaggcc aaagtggaa
751 gatggagttc ttctggacaa ttttaaagcc gaatgatgca atcaacttcg
801 agagtaatgg aaatttcatt gctccagaat atgcatacaa aattgtcaag
851 aaaggggact caacaattat gaaaagtga ttggaatatg gtaactgcaa
901 caccaagtgt caaactccaa tgggggcat aaactctagc atgccattcc
951 acaatataca ccctctcacc attggggaat gcccacaata tgtgaaatca
1001 aacagattag tccttgcgac tgggctcaga aatagccctc aaagagagac
1051 tcgaggatta tttggagcta tagcaggttt tatagagga ggatggcagg
1101 gaatggtaga tggttggtat gggtagcacc atagcaatga gcaggggagt
1151 gggtagctg cagacaaaga atccactcaa aaggcaatag atggagtcac
1201 caataaggtc aactcgatca ttgacaaaat gaactctag tttgaggccg
1251 ttggaaggga atttaacaac ttagaaagga gaatagagaa tttaaacaag
1301 aagatggaag acgggttcct agatgtctgg acttataatg ctgaacttct
1351 ggttctcatg gaaaatgaga gaactctaga ctttcatgac tcaaagtca
1401 agaaccctta cgacaaggtc cgactacagc ttagggataa tgcaaaggag
1451 ctgggtaacg gttgtttcga gttctatcat aaatgtgata atgaatgtat
1501 ggaaagtgtg agaaatggaa cgtatgacta cccgcagtat tcagaagaag
1551 cgagactaaa aagagaggaa ataagtggag taaaattgga atcaatagga
1601 atttaccaaa tactgtcaat ttattctaca gtggcgagtt ccctagcact
1651 ggcaatcatg gtagctggtc tatccttatg gatgtgctcc aatgggtcgt
1701 tacaatgcag aatttgcatt taaatttggt agttcagatt gtagttaaaa
1751 acacccttgt ttctact

```

Secuencia de aminoácidos de A/Vietnam/1203/04 H5 (SEC ID N° 11) Longitud de la molécula completa: 564 aa

```

1 mektivllfai vslvksdqic igyhannste qvdtimeknv tvthaqdile
51 kkhngklcdl dgvkplilrd csvagwllgn pmcdefinvp ewsyivekan
5 101 pvndlcypgd fndyeelkhl lsrinhfeki qiipksswss heaslgvssa

151 cpyqgkssff rnvwlikkn styptikrsy nntnqedllv lwgihhpnda
201 aeqtklyqnp ttyisvgtst lnqrlvpria trskvngqsg rmeffwtilk
251 pndainfesn gnfiapayay kivkkgsti mkseleygnc ntkcqtpmga
301 inssmpfhni hpltigecpk yvksnrlvla tglrnspqre trglfgaiag
351 fieggwqgmvg dwygyhhsn eqsgyaadk estqkaidgv tnkvnsiidk
401 mntqfeavgr efnllerrie nlnkkmedgf ldvwtynael lvlmenertl
451 dfhdnsvknk ydkvrlqlrd nakelngcf efyhkcdnec mesvrngtyd
501 ypyyseearl kreeisgvkl esigiyqils iystvassla laimvagls
551 wmcnsgslqc rici

```

Secuencia de nucleótidos de A/Vietnam/1203/04 N1 (SEC ID N° 2) Longitud de la molécula completa: 1398 nt

```

1 agcaaaagca ggagttcaaa atgaatccaa atcagaagat aataaccatc
51 ggggtcaatct gtatggtaac tggaatagtt agcttaatgt tacaaattgg
101 gaacatgata tcaatatggg tcagtcattc aattcacaca ggggaatcaac
151 accaatctga accaatcagc aataactaatt ttcttactga gaaagctgtg
201 gcttcagtaa aattagcggg caattcatct ctttgcccca ttaacggatg
251 ggctgtatac agtaaggaca acagtataag gatcggttcc aagggggatg
301 tgtttgttat aagagagccg ttcattcat gctcccaactt ggaatgcaga
351 actttctttt tgactcaggg agccttgctg aatgacaagc actccaatgg
401 gactgtcaaa gacagaagcc ctcacagaac attaatgagt tgcctgtgg
451 gtgaggctcc ctccccatat aactcaagg ttagtctgtg tgcttggtea
501 gcaagtgtt gccatgatgg caccagttgg ttgacgattg gaatttctgg
551 cccagacaat ggggctgtgg ctgtattgaa atacaatggc ataataacag
601 acactatcaa gagttggagg aacaacatac tgagaactca agagtctgaa
651 tgtgcatgtg taaatggctc ttgctttact gtaatgactg acggaccaag
701 taatggtcag gcatcacata agatcttcaa aatggaaaaa gggaaagtgg
751 ttaaatacgt cgaattggat gctcctaatt atcactatga ggaatgctcc
801 tgttatccta atgccggaga aatcacatgt gtgtgcaggg ataattggca
851 tggtctaaat cggccatggg tatctttcaa tcaaaatttg gagtatcaaa
901 taggatatat atgcagtggg gttttcggag acaatccacg cccaatgat
951 ggaacaggta gttgtggtcc ggtgtcctct aacggggcat atggggtaaa
1001 aggggtttca tttaaatacg gcaatggtgt ctggatcggg agaaccaaaa
1051 gcactaatcc caggagcggc tttgaaatga tttgggatcc aaatgggtgg
1101 actgaaacgg acagtagctt ttcagtgaag caagatatcg tagcaataac
1151 tgattggtea ggatatagcg ggagttttgt ccagcatcca gaactgacag
1201 gactagattg cataagacct tgtttctggg ttgagttgat cagagggcgg
1251 cccaaagaga gcacaatttg gactagtggg agcagcatat ctttttgtgg
1301 tgtaaatagt gacactgtgg gttggtcttg gccagacggg gctgagttgc
1351 cattcaccat tgacaagtag tttgttcaaa aaactccttg tttctact

```

Secuencia de aminoácidos de A/Vietnam/1203/04 N1 (SEC ID N° 12) Longitud de la molécula completa: 449 aa

```

1 mnpnqkiiti gsicmvtgiv slmlqignmi siwvshsiht gnqhqsepis
51 ntnfltekav asvklagnss lcpingwavy skdnsirigs kgdvvfirep
101 fiscshlecr tffltqgall ndkhsngtvk drsphrtlms cpvgeapspy
5 151 nsrfesvaws asachdgtsw ltigisgpdn gavavlkyng iitdtikswr

201 nnilrtqese cacvngscft vmtdgpsngg ashkifmek gkvvksveld
251 apnyhyeecs cypnageitc vcrdnwhgsn rpwvsfnqnl eyqigyicsg
301 vfgdnprpnd gtgscgpvss ngaygvkgfs fkygngvwig rtkstnsrsg
351 femiwdpngw tetdssfsvk qdivaitdws gysgsfvqhp eltglldcirp
401 cfwvelirgr pkestiwtsg ssisfcgvns dtvgwswpdg aelpftidk

```

A/Hong Kong/213/03

10

Secuencia de nucleótidos de A/Hong Kong/213/03 H5 (SEC ID N° 3) Longitud de la molécula completa: 1767 nt

```

1 agcaaaagca ggggttcaat ctgtcaaaat ggagaaaata gtgcttcttt
51 ttgcaatagt cagtcttggt aaaagtgatc agatttgcat tggttaccat
101 gcaaacaact cgacagagca ggttgacaca ataatggaaa agaacgttac
151 tgttacacat gccaagaca tactggaaaa gacacacaac gggaagctct
201 gcgatctaga tggagtgaag cctctaattt tgagagattg tagtgtagct
251 ggatggctcc tcggaaaccc aatgtgtgac gaattcatca atgtgccgga
301 atggtcttac atagtggaga aggccaatcc agccaatgac ctctgttacc
351 caggggattt caacgactat gaagaattga aacacctatt gagcagaata
401 aaccatthtt agaaaattca gatcatcccc aaaaatttctt ggtccagtca
451 tgaagcctca ttaggggtga gctcagcatg tccataccaa ggaaagtcct
501 cctttttcag gaatgtggta tggcttatca aaaagaacaa tgcataccca
551 acaataaaga ggagctacaa taataccaac caagaagatc ttttgggtatt
601 gtgggggatt caccatccta atgatgcggc agagcagact aggtctctatc
651 aaaacccaac cacctacatt tccgttggga catcaacact aaaccagaga
701 ttggtaccaa aaatagctac tagatccaaa gtaaacgggc aaaatggaag
751 gatggagttc ttctggacaa ttttaaaacc gaatgatgca atcaacttcg
801 agagcaatgg aaatttcatt gctccagaat atgcatacaa aattgtcaag
851 aaaggggact cagcaattat gaaaagtga tttggaatatg gtaactgcaa
901 caccaagtgt caaactccaa tgggggagat aaactctagt atgccattcc
951 acaatataca ccctctcacc atcggggaat gcccacaaata tgtgaaatca
1001 aacagattag tccttgcgac tgggctcaga aatagccctc aaagagagac
1051 tcgaggatta tttggagcta tagcagggtt tatagaggga ggatggcagg
1101 gaatggtaga tggttggtat ggggtaccacc atagcaatga gcaggggagt
1151 gggtagctg cagacaaaga atccactcaa aaggcaatag atggagtcac
1201 caataaggtc aactcgatca ttgacaaaat gaacactcag tttgaggccg
1251 ttggaaggga atttaataac ttagaaagga gaatagagaa tttaaacaag
1301 aagatggaag acggattcct agatgtctgg acttataatg ctgaacttct
1351 ggttctcatg gaaaatgaga gaactctaga ctttcatgac tcaaattgtca
1401 agaaccctta cgacaaggtc cgactacagc ttagggataa tgcaaaggag
1451 ctgggtaacg gttgtttcga gttctatcac aaatgtgata atgaatgtat
1501 ggaaagtgtg agaaacggaa cgtatgacta cccgcagtat tcagaagaag
1551 caagactaaa aagagaggaa ataagtggag taaaattgga gtcaatagga
1601 acttaccaa tactgtcaat ttattctaca gtggcgagtt ccctagcact
1651 ggcaatcatg gtagctggtc tatctttatg gatgtgctcc aatgggtcgt
1701 tacaatgcag aatttgcatt taaatttgtg agttcagatt gtagttaaaa
1751 acacccttgt ttctact

```

Secuencia de aminoácidos de A/Hong Kong/213/03 H5 (SEC ID N° 13) Longitud de la molécula completa: 564 aa

```

1 mektivllfai vslvksdqic igyhannste qvdtimeknv tvthaqdile
51 kthngklcdl dgvkplilrd csvagwllgn pmcdefinvp ewsyivekan
101 pandlcypgd fndyeelkhl lsrinhfeki qiipknswws heaslgvssa
151 cpyqgkssff rnvwlikkn nayptikrsy nntngedllv lwgihhpnda
201 aeqtrlyqnp ttyisvgtst lnqrlvpkia trskvngqng rmeffwtilk
251 pndainfesn gnfiapeyay kivkkgsai mkseleygnc ntkcqtpmga
301 inssmpfhni hpltigecpk yvksnrlvla tglrnspqre trglfgaiag
351 fieggwgqmv dgwygyhhsn eggsgyaadk estqkaidgv tnkvnsiidk
401 mntqfeavgr efnllerrie nlnkkmedgf ldvwtynael lvlmenertl
451 dfhdsnvknl ydkvrlqlrd nakelngngcf efyhkcdnec mesvrngtyd
501 ypyyseearl kreeisgvkl esigtyqils iystvassla laimvaglsl
551 wmcsmgslqc rici

```

Secuencia de nucleótidos de A/Hong Kong/213/03 N1 (SEC ID N° 4) Longitud de la molécula completa: 1458 nt

```

1 agcaaaaagca ggagttcaaa atgaatccaa atcagaagat aacaaccatt
51 ggatcaatct gtatggtaat tggaatagtt agcttgatgt tacaaattgg
101 gaacataatc tcaatatggg ttagtcattc aattcaaaca gggaaatcaac
151 accaggctga accatgcaat caaagcatta ttacttatga aaacaacacc
201 tgggtaaacc agacatatgt caacatcagc aataccaatt ttcttactga
251 gaaagctgtg gcttcagtaa cattagcggg caattcatct ctttgcccca
301 ttagtgatg ggctgtatac agtaaggaca acggtataag aatcggttcc
351 aagggggatg tgtttggtat aagagagccg ttcattctcat gctcccactt
401 ggaatgcaga actttctttt tgactcaggg agccttgctg aatgacaagc
451 attctaattg gaccgtcaaa gacagaagcc ctacacagaac attaatgagt
501 tgtcccgtgg gtgaggtccc ttccccatac aactcgaggt ttgagtctgt
551 tgcttggtcg gcaagtgctt gtcattgatg cactagttgg ttgacaattg
601 gaatttcttg cccagacaat ggggctgtgg ctgtattgaa atacaatggc
651 ataataacag acactatcaa gagttggagg aacaacataa tgagaactca
701 agagtctgaa tgtgcatgtg taaatggctc ttgctttact gttatgactg
751 atggaccaag taatgggcag gcttcataca aaatcttcag aatagaaaaa
801 gggaaagtag ttaaatacag cgaattaaat gcccctaatt atcactatga
851 ggagtgtccc tggtatcctg atgctggaga aatcacatgt gtgtgcaggg
901 ataactggca tggctcaaat cggccatggg tatctttcaa tcaaaatttg
951 gagtatcgaa taggatatat atgcagtggg gttttcggag acaatccacg
1001 cccaatgat gggacaggca gttgtggtcc ggtgtcccct aaaggggcat
1051 atggaataaa agggttctca tttaaatac gcaatggtgt ttggatcggg
1101 agaaccaaaa gcactaattc caggagcggc tttgaaatga tttgggatcc
1151 aaatggatgg actggtacgg acagtaattt ttcagtaaag caagatattg
1201 tagctataac cgattggtca ggatatagcg ggagttttgt ccagcatcca
1251 gaactgacag gattagattg cataagacct tgtttctggg ttgagctaat
1301 cagagggcgg cccaaagaga gcacaatttg gactagtggg agcagcatat
1351 ccttttgtgg tgtaaatagt gacactgtgg gttggtcttg gccagacggt
1401 gctgagttgc cattcaccat tgacaagtag tttgttcaaa aaactccttg
1451 tttctact

```

5

Secuencia de aminoácidos de A/Hong Kong/213/03 N1 (SEC ID N° 14)

Longitud de la molécula completa: 469 aa

10

```

1 mnpnqkitti gsicmvigiv slmlqignii siwvshsiqt gnqhqaepcn
51 qsiityennt wvnqtyvnis ntnfltekav asvtlagness lcpisgwavy
101 skdngirigs kgdvvfirep fiscshlecr tffltqgall ndkhsngtvk
151 drsphrtlms cpvgeapsy nsrfesvaws asachdgtsw ltigisgpdn
201 gavavlk yng iitdtikswr nnimrtqese cacvngscft vmtdgpsngq
251 asykifriek gkvvksaeln apnyhyeecs cypdageitc vcrdnwhgsn
301 rpwvsfnqnl eyrigyicsg vfgdnprpnd gtgscgpvsp kgaygikgfs
351 fkygngvwig rtkstnsrsg femiwdpngw tgtdsnfsvk qdivaitdws
401 gysgsfvqhp eltglcdirp cfwvelirgr pkestiwtsg ssisfcgvns
451 dtvgwswpdg aelpftidk

```

A/Hong Kong/491/97 (HA) + A/Hong Kong/486/97 (NA)

5 Secuencia de nucleótidos de A/Hong Kong/491/97 H5 (SEC ID N° 5) Longitud de la molécula completa: 1767 nt

```

1 agcaaaagca ggggtataat ctgtcaaaat ggagaaaata gtgcttcttc
51 ttgcaacagt cagccttggt aaaagtgacc agatttgcac ttggttaccat
101 gcaaaacaact cgacagagca agttgacaca ataattggaaa agaattgttac
151 tgttacacat gcccaagaca tactggaaag gacacacaac gggaagctct
201 gcgatctaaa tggagtgaag cctctgattt tgagggtatt tagtgtagct
251 ggatggctcc tcggaaaccc tatgtgtgac gaattcatca atgtgccgga
301 atggtcttac atagtggaga aggccagtcc agccaatgac ctctgttatc
351 cagggaattt caacgactat gaagaactga aacacctatt gagcagaata
401 aaccattttg agaaaattca gataatcccc aaaagttctt ggtccaatca
451 tgatgcctca tcaggggtga gctcagcatg tccatacctt gggaggtcct
501 cctttttcag aaatgtggta tggcttatca aaaagaacag tagctacca
551 acaataaaga ggagctacaa taataccaac caagaagatc ttttgggtact
601 gtgggggatt caccatccta atgatgcggc agagcagaca aggtctctac
651 aaaacccaac cacctacatt tccgttgga catcaacact gaaccagaga
701 ttggttccag aaatagctac tagacccaaa gtaaacgggc aaagtggaa
751 aatggagttc ttctggacaa ttttaaagcc gaatgatgcc atcaatttcg
801 agagtaatgg aaatttcatt gctccagaat atgcatacaa aattgtcaag
851 aaaggggact caacaattat gaaaagtga ttggaatatg gtaactgcaa
901 caccaagtgt caaactccaa tgggggcaat aaactctagt atgccattcc
951 acaacataca cccctcacc atcggggaat gcccacaaata tgtgaaatca
1001 aacagattag tccttgcaac tggactcaga aatacccctc aacgagagac
1051 gcgaggacta tttggagcta tagcagggtt tatagaggga ggatggcagg
1101 gaatggtaga tgggtggtat ggggtaccac atagcaatga gcaggggagt
1151 ggatacgtg cagaccaaga atccacacaa aaggcaatag atggagtcac
1201 caataagggtc aactcgatca ttaacaaaat gaacactcag tttgaggccg
1251 ttggaaggga atttaataac ttggaaagga ggatagagaa tttaaacaag
1301 aaaatggaag acggattcct agatgtctgg acttacaatg ccgaacttct
1351 ggttctcatg gaaaatgaga gaactctaga ctttcatgac tcaaattgtca
1401 agaaccttta cgacaagggt cgactacagc ttagggataa tgcaaaggag
1451 ctgggtaatg gttgtttcga attctatcac aaatgtgata acgaatgtat
1501 ggaaagtgtg aaaaacggaa cgtatgacta cccgcagtat tcagaagaag
1551 caagactaaa cagagaggaa ataagtggag taaaattgga atcaatggga
1601 acttaccaaa tactgtcaat ttattcaaca gtggcgagtt ccctagcact

```

1651 ggcaatcatg gtagctggtc tatctttatg gatgtgctcc aatggatcgt  
 1701 tacaatgcag aatttgcatt taaatttgtg agttcagatt gtagttaaaa  
 1751 acacccttgt ttctact

Secuencia de aminoácidos de A/Hong Kong/491/97 H5 (SEC ID N° 15) Longitud de la molécula completa: 564 aa

1 mektivlllat vslvksdqic igyhannste qvdtimeknv tvthaqdile  
 51 rthngklcdl ngvkplilrd csvagwllgn pmcdefinvp ewsyivekas  
 101 pandlcypgn fndyeelkhl lsrinhfeki qiipksswsn hdassgvssa  
 151 cpylgrssff rnvwvlikn ssyptikrsy nntnqedllv lwgihhpnda  
 201 aeqtrlyqnp ttyisvgtst lnqrlvpeia trpkvngqsg rmeffwtilk  
 251 pndainfesn gnfiapeyay kivkkgdsti mkseleygnc ntkcqtpmga  
 301 inssmpfhni hpltigecpk yvksnrlvla tglrntpqr trglfgaiag  
 351 fieggwqgmvg dgwygyhhsn eqgsgyaadq estqkaidgv tnkvnsiink  
 401 mntqfeavgr efnnlerrie nlnkkmedgf ldvwtyna el lvmenertl  
 451 dfhdsnvknl ydkvrlqlrd nakelngncf efyhkcdnec mesvkngtyd  
 501 ypqyseearl nreeisgvkl esmgtyqils iystvassla laimvagls  
 551 wmcnsngslqc rici

5

Secuencia de nucleótidos de A/Hong Kong/486/97 N1 (SEC ID N° 6) Longitud de la molécula completa: 1401 nt

1 agcaaaaagca ggagttttaa atgaatccaa atcagaagat aataaccatt  
 51 ggatcaatct gcatggtagt tgggataatc agcttgatgt tacaatttgg  
 101 aaacacaata tcagtatggg tcagccacat aattaaaact tggcacccaa  
 151 accagcctga accatgcaac caaagcatca atttttacac tgagcaggct  
 201 gcagcttcag tgacattagc gggcaattcc tctctctgcc ctatttagtgg  
 251 atgggctata tacagcaagg acaatagtat aagaattggg tccaaagggg  
 301 atgtgtttgt tataagagaa ccattcatct catgctccca tttggaatgc  
 351 agaacctttt tcttgacca aggagcccta ttgaatgaca agcattctaa  
 401 tgggaccgtc aaagacagga gcccttatag aactttaatg agctgtcctg  
 451 ttggtgaggc cccttcccca tacaactcaa ggtttgagtc tgttgcttgg  
 501 tcagcaagtg cttgccatga tggcattagt tggctaaca ttggaatttc  
 551 cgggtccgat aatggggctg tggctgtgtt gaaatacaat ggcataataa  
 601 cagacaccat caagagttgg aggaacaaca cactgaggac gcaagagtct  
 651 gaatgtgcat gtgtgaatgg ttcttggttt actgtaatga cagatggacc  
 701 gagtaatgaa caggcctcat acaagatttt caagatagaa aaggggaggg  
 751 tagtcaaadc agttgagttg aacgccccta attatcatta cgaggaatgc  
 801 tcttggtatc ctgatgctgg cgaaatcaca tgtgtgtgca gggataattg  
 851 gcatggctcg aaccgaccat ggggtgtctt caatcagaat ctggagtatc  
 901 aaataggata tatatgcagt ggggttttcg gagacagtcc acgcccacat  
 951 gatgggacag gcagttgtgg tccagtgtct cttaacggag cgtatggagt  
 1001 aaaaggggtt tcattttaat acggcaatgg tgtttggatc gggagaacca  
 1051 aaagcactag ttccaggagc ggttttgaaa tgatttggga tccaaatggg  
 1101 tggaccgaaa cagacagtag cttctcgttg aagcaagaca tcatagcgat  
 1151 aactgattgg tcaggataga gcgggagttt tattcaacat ccagaactga  
 1201 caggattaaa ttgcatgaga ccttgcttct gggttgaact aatcagaggg  
 1251 aggcccaaag agaaaacaat ctggactagt gggagcagta tatctttctg  
 1301 tgggtgtaaat agtgacactg tgggttggtc ttggccagac ggtgctgagt

10

1351 tgccatacac cattgacaag tagtttggtc aaaaaactcc ttgtttctac  
1401 t

Secuencia de aminoácidos de A/Hong Kong/486/97 N1 (SEC ID Nº 16) Longitud de la molécula completa: 450 aa

1 mnpnqkiiti gsicmvvgii slmlqignti svwvshiikt whpnqpepcn  
51 qsinfyteqa aasvtlagns slcpisgwai yskdnsirig skgdivfvire  
101 pfiscshlec rtffltggal lndkhsngtv kdrspyrtlm scpvgeapsp  
151 ynsrfesvaw sasachdgis wltigisgpd ngavavlkyn giitdtiksw  
201 rnntlirtqes ecacvngscf tvmtdgpsne qasykifkie kgrvvksvel  
251 napnyhyeec scypdageit cvcrdnwhgs nrpwwsfngn leyqigyics  
301 gvfgdsprpn dgtgscgpvs lngaygvkgf sfkygngvwi grtkstssrs  
351 gfemiwdpng wtetdssfsi kqdiiaitdw sgysgsfiqh peltglncmr  
5 401 pcfwvelirg rpkektiws gssisfcgvn sdtvgwswpd gaelpytidk

A/Hong Kong/491/97 (Ser211) (HA) + A/Hong Kong/486/97 (NA)

Secuencia de nucleótidos de A/Hong Kong/491/97 (Ser211) H5 (SEC ID Nº 7) Longitud de la molécula completa:  
10 1767 nt

1 agcaaaagca ggggtataat ctgtcaaaat ggagaaaata gtgcttcttc  
51 ttgcaacagt cagccttggt aaaagtgacc agatttgcat tggttaccat  
101 gcaaacact cgacagagca agttgacaca ataattggaaa agaattgttac  
151 tgttacacat gccaagaca tactggaaaag gacacacaac ggaagctct  
201 gcgatctaaa tggagtgaag cctctgattt tgagggattg tagtgtagct  
251 ggatggctcc tcggaaaccc tatgtgtgac gaattcatca atgtgccgga  
301 atgggtcttac atagtggaga aggccagtcc agccaatgac ctctgttatac  
351 caggggaattt caacgactat gaagaactga aacacctatt gagcagaata  
401 aaccattttg agaaaattca gataatcccc aaaagttctt ggtccaatca  
451 tgatgcctca tcaggggtga gctcagcatg tccatacctt gggaggctct  
501 cctttttcag aaatgtggta tggcttatca aaaagaacag tagctacca  
551 acaataaaga ggagctacaa taataccaac caagaagatc ttttggtact  
601 gtgggggatt caccatccta atgatgcggc agagcagaca aggctctatc  
651 aaaacccaac cacctacatt tccgttgga catcaacact gaaccagaga  
701 ttgggtttcag aaatagctac tagacccaaa gttaaaggggc aaagtggag  
751 aatggagttc ttctggacaa ttttaaagcc gaatgatgcc atcaatttcg  
801 agagtaatgg aaatttcatt gctccagaat atgcatacaa aattgtcaag  
851 aaaggggact caacaattat gaaaagtga ttggaatatg gtaactgcaa  
901 caccaagtgt caaactccaa tgggggcaat aaactctagt atgccattcc  
951 acaacatata cccctcacc atcggggaat gcccacaaata tgtgaaatca  
1001 aacagattag tccttgcaac tggactcaga aatacccctc aacgagagac  
1051 gcgaggacta tttggagcta tagcaggttt tatagaggga ggatggcagg  
1101 gaatggtaga tggttggtat gggtagacc atagcaatga gcaggggagt  
1151 ggatacgtg cagaccaaga atccacacaa aaggcaatag atggagtcac  
1201 caataagggtc aactcgatca ttaacaaaat gaacactcag tttgaggccg  
1251 ttggaaggga atttaataac ttggaaagga ggatagagaa tttaaacaag  
1301 aaaatggaag acggattcct agatgtctgg acttacaatg ccgaacttct

```

1351 ggttctcatg gaaaatgaga gaactctaga ctttcatgac tcaaattgtca
1401 agaaccttta cgacaagggtc cgactacagc ttagggataa tgcaaaggag
1451 ctgggtaatg gttgtttcga attctatcac aaatgtgata acgaatgtat
1501 ggaaagtgtg aaaaacggaa cgtatgacta cccgcagtat tcagaagaag
1551 caagactaaa cagagaggaa ataagtggag taaaattgga atcaatggga
1601 acttaccaaa tactgtcaat ttattcaaca gtggcgagtt ccctagcact
1651 ggcaatcatg gtagctgggtc tatctttatg gatgtgctcc aatggatcgt
1701 tacaatgcag aatttgcatt taaatttgtg agttcagatt gtagttaaaa
1751 acacccttgt ttctact

```

Secuencia de aminoácidos de A/Hong Kong/491/97 (Ser211) H5 (SEC ID N° 17) Longitud de la molécula completa: 564 aa

5

```

1 mekiivlllat vslvksdqic igyhannste qvdttimeknv tvthaqdile
51 rthngklcdl ngvkplilrd csvagwllgn pmcdefinvp ewsyivekas
101 pandlcypgn fndyeelkhl lsrinhfeki qiipksswsn hdassgvssa
151 cpylgrssff rnvvwlikkn ssyptikrsy nntnqedllv lwgihhpnda
201 aeqtrlyqnp ttyisvgtst lnqrlvseia trpkvngqsg rmeffwtilk
251 pndainfesn gnfiapeyay kivkkgdsti mkseleygnc ntkcqtpmga
301 inssmpfhni hpltigecpk yvksnrlvla tglrntpqre trglfgaiag
351 fieggwqgmv dgwygyhhsn eggsgyaadq estqkaidgv tnkvnsiink
401 mntqfeavgr efnnlerrrie nlnkkmedgf ldvwtynael lvlmenertl
451 dfhdsnvknl ydkvrlqlrd nakelngnfcf efyhkcdnec mesvkngtyd
501 ypqyseearl nreeisgvkl esmgtyqils iystvassla laimvagls1
551 wmcnsngslqc rici

```

Secuencia de nucleótidos de A/Hong Kong/486/97 N1 (SEC ID N° 8) Longitud de la molécula completa: 1401 nt

```

1 agcaaaaagca ggagttttaa atgaatccaa atcagaagat aataaccatt
51 ggatcaatct gcatggtagt tgggataatc agcttgatgt tacaatttgg
101 aaacacaata tcagtatggg tcagccacat aattaaaact tggcacccaa
151 accagcctga accatgcaac caaagcatca atttttacac tgagcaggct
201 gcagcttcag tgacattagc gggcaattcc tctctctgcc ctattagtgg
251 atgggctata tacagcaagg acaatagtat aagaattggt tccaaagggg
301 atgtgtttgt tataagagaa ccattcatct catgctccca tttggaatgc
351 agaacctttt tcttgacca aggagcccta ttgaatgaca agcattctaa
401 tgggaccgtc aaagacagga gccctatag aactttaatg agctgtcctg
451 ttggtgaggc cccttcccca tacaactcaa ggtttgagtc tgttgcttgg
501 tcagcaagtg cttgccatga tggcattagt tggctaacaa ttggaatttc
551 cggtcgggat aatggggctg tggctgtgtt gaaatacaat ggcataataa
601 cagacaccat caagagttgg aggaacaaca cactgaggac gcaagagtct
651 gaatgtgcat gtgtgaatgg ttcttgtttt actgtaatga cagatggacc
701 gagtaatgaa caggcctcat acaagatttt caagatagaa aaggggaggg
751 tagtcaaate agttgagttg aacgccccta attatcatta cgaggaatgc
801 tcctgttate ctgatgctgg cgaaatcaca tgtgtgtgca gggataattg
851 gcatggctcg aaccgacat ggggtgtctt caatcagaat ctggagtatc
901 aaataggata tatatgcagt ggggttttcg gagacagtcc acgccccaat
951 gatgggacag gcagttgtgg tccagtgtct cttaacggag cgtatggagt
1001 aaaagggttt tcattttaat acggcaatgg tgtttggatc gggagaacca

```

10



1051 aaagcactag ttccaggagc ggttttgaaa tgatttgga tccaaatggg  
 1101 tggaccgaaa cagacagtag cttctcggtg aagcaagaca tcatagcgat  
 1151 aactgattgg tcaggataca gcgggagttt tattcaacat ccagaactga  
 1201 caggattaaa ttgcatgaga ccttgcttct ggggtgaact aatcagaggg  
 1251 aggcccaaag agaaaacaat ctggactagt gggagcagta tatctttctg  
 1301 tgggtgtaa atgtgacactg tgggttggtc ttggccagac ggtgctgagt  
 1351 tgccatacac cattgacaag tagtttggtc aaaaaactcc ttgtttctac  
 1401 t

Secuencia de aminoácidos de A/Hong Kong/486/97 N1 (SEC ID N° 18) Longitud de la molécula completa: 450 aa

1 mnpnqkiiti gsicmvvgii slmlqignti svwvshiikt whpnqpepcn  
 51 qsinfytega aasvtlagns slcpisgwai yskdnsirig skgdvfvire  
 101 pfiscshlec rtffltqgal lndkhsngtv kdrspyrtlm scpvgeapsp  
 151 ynsrfesvaw sasachdgis wltigisgpd ngavavlkyn giitdtiksw  
 201 rnntlrtltes ecacvngscf tvmtdgpsne qasykifkie kgrvvksvel  
 251 napnyhyeec scypdageit cvcrdnwhgs nrpwwsfngn leyqigyics  
 301 gvfgdsprpn dgtgscgpvs lngaygvkgf sfkyngvwi grtkstssrs  
 351 gfemiwdpng wtetdssfsf kqdiiaitdw sgysgsfiqh peltglncmr  
 5 401 pcfwvelirg rpkektiws gssisfcgvn sdtvgwswpd gaelpytidk

ca A/ck/Hong Kong/G9/97

Secuencia de nucleótidos de ca A/ck/Hong Kong/G9/97 (SEC ID N° 9) Longitud de la molécula completa: 1690 pb

1 ttaaccactc aagatggaag caataccact aataactata ctactagtag  
 51 taacagcaag caatgcagac aaaatctgca tcggctacca atcaacaaac  
 101 tccacagaaa ccgtagacac gctaacagaa aacaatgttc ctgtgacaca  
 151 tgccaaagaa ttgctccaca cagagcacia tgggatgctg tgtgcaacaa  
 201 atctgggacg tcctcttatt ctgacacatt gcaccattga aggactgac  
 251 tatggcaacc cttcttgtga tctactgttg ggaggaagag aatggtccta  
 301 catcgctgaa agaccatcgg ctgttaatgg aatgtgttac cccgggaatg  
 351 tagaaaacct agaggaacta aggtcatttt ttagttctgc tagttcctac  
 401 caaagaatcc agatctttcc agacacaatc tggaatgtgt cttacagtgg  
 451 aacaagcaaa gcatgttcag attcattcta caggagcatg agatggttga  
 501 ctcaaaagaa caacgcttac cctattcaag acgccaata cacaaataat  
 551 agaggaaaga gcattctttt catgtggggc ataatcacc cacctaccga  
 601 tactgcacag acaaatctgt acacaaggac tgacacaaca acaagtgtgg  
 651 caacagaaga tataaatagg accttcaaac cagtgatagg gccaaggccc  
 701 cttgtcaatg gtctgcaggg aagaattgat tattattggg cggatttgaa  
 751 accaggtcag acattgagag taagatccaa tgggaatcta atcgctccat  
 801 ggtatgggca cattctttca ggagagagcc acggaagaat cctgaagact  
 851 gattttaaaca gtggtagctg tgtagtgcac tgtcaaacag aaagaggtgg  
 901 cttaaatact actttgccat tccacaatgt cagtaaataat gcatttgga  
 951 actgcccaca atatgttgga gtaaagagtc tcaaactggc agttggtctg  
 1001 aggaatgtgc ctgctagatc aagtagagga ctatttgggg ccatagctgg  
 1051 attcatagag ggaggttggt cagggtggt cgctgggttg tatgggttcc  
 1101 agcattcaaa tgatcaaggg gttggtatag ctgcagatag agactcaact

```

1151 caaagggcaa ttgacaaaat aacgtccaaa gtgaataata tagtcgataa
1201 aatgaacaag caatatgaaa ttattgatca tgaattcagc gaggttgaaa
1251 atagactcaa tatgatcaat aataagattg atgaccagat acaagacata
1301 tgggcatata acgctgaatt gctagtgtg cttgaaaacc agaaaacact
1351 cgatgagcat gatgcgaatg taaacaatct atataacaaa gtgaagaggg
1401 cactgggttc caatgcaatg gaagatggga aaggatgttt cgagctatac
1451 cataaatgtg atgatcagt catggagaca attcggaacg ggacctataa
1501 caggaggaag tataaagagg aatcaagact agaaagacag aaaatagaag
1551 gggtaagct ggaatctgaa ggaacttaca aaatcctcac catttattcg
1601 actgtcgct catctctgt gattgcaatg gggtttgctg ccttcttgtt
1651 ctggggccatg tccaatggat cttgcagatg caacatttga

```

Secuencia de aminoácidos de ca A/ck/Hong Kong/G9/97 H9 (SEC ID N° 19) Longitud de la molécula completa: 558 aa

5

```

1  meaiplitil lvvtasnadk icigyqstns tetvdtlten nvpvthakel
51  lhtehngmlc atnlgrplil dtctiegly gnpscdlllg grewsyiver
101 psavngmcyp gnvenleelr sffssassyq riqifpdtiw nvsysgtska
151 csdsfyrsmr wltqknnayp iqdaqytnnr gksilfmwgi nhpptdtaqt
201 nlytrtdttt svatedinrt fkpvigprpl vnglqgridy ywsvlkgpgt
251 lrvrsngnli apwyghilsg eshgrilkt lnsqscvvqc qtergglnnt
301 lpfhvnvskya fgncpkyvgv kslklavglr nvparssrgl fgaiagfieg
351 gwsglvagwy gfqhsndqgv giaadrdest raidkitskv nnivdkmnkq
401 yeiidhefse venrlminn kiddqiqdiw aynaellvll enqktldehd
451 anvnnlynkv kralgsname dgkgcfelyh kcddqcmeti rngtynrrky
501 keesrlerqk iegvkleseg tykiltiyst vasslviamg faaflfwams
551 ngscrcni

```

Secuencia de nucleótidos de ca A/ck/Hong Kong/G9/97 N2 (SEC ID N° 10) Longitud de la molécula completa: 1428 pb

10

```

1  aatgaatcc aatcagaag ataatagcaa ttggctctgt ttctctaact
51  attgcgacaa tatgcctcct catgcagatt gctatcttag caacgactat
101 gacactacat ttcaagcaga atgaatgcat caactcctcg aataatcaag
151 tagtgccatg tgaaccaatc ataatagaaa ggaacataac agagatagtg
201 catttgaata gtactacctt agagaaggaa atttgtccta aagtagcaga
251 ctacaggaat tgggtcaaac cacaatgtca aatcacaggg ttcgctcctt
301 tctccaagga caattcaatt aggctctccg caggtggaga tatttgggtg
351 acaagagAAC cttatgtatc gtgcggctct ggtaaatgtt atcaatttgc
401 acttggggcag ggaaccactt tggagaacaa acactcaaac ggcacagcac
451 atgatagaac tcctcataga acccttttaa tgaatgagtt ggggtgttccg
501 tttcatttgg caaccaaaca agtgtgcata gcatgggtcca gctcaagctg
551 ccatgatggg aaagcatggt tacatgtttg tgtcactggg gatgatagaa
601 atgcaacggc tagcatcatt tatgatggga tacttgttga cagtattggt
651 tcatggtcta aaaacatcct cagaactcag gagtcagaat gcgtttgcat
701 caatggaacc tgtgcagtag taatgactga tggaaagtga tcaggaaggg
751 ctgacactag aatactatct attagagagg ggaaaattgc acacattagc
801 ccattgtcag gaagtgtcga gcatgtggag gaatgtcctt gttacccccg
851 atatccagaa gttagatgtg tttgcagaga caattggaag ggatccaata
901 ggcccgttct atatataaat atggcaaatt atagtattga ttccagttat

```

951 gtgtgctcag gacttggttg cgacacacca agaaatgatg ataggtctag  
 1001 cagcagcaac tgcagagatc ctaataacga gagaggggccc ccaggagtaa  
 1051 aagggtgggc ctttgacaat ggaaatgaca tttggatggg aagaacaatc  
 1101 aaaaaggatt cgcgctcagg ttatgagact ttcagggtca ttggtggttg  
 1151 gaccactgct aattccaagt cacagataaa tagacaagtc atagttgaca  
 1201 gtgataactc gtctgggtat tctgggtatct tctctgttga aggcaaaagc  
 1251 tgcatacaaca ggtgttttta cgtggagttg ataagaggaa gaccaaagga  
 1301 gactaggggtg tgggtggactt caaatagcat cattgtattt tgtggaactt  
 1351 caggtaccta tgggaacaggc tcatggcctg atggggcgaa tatcaatttc  
 1401 atgcctatat aagctttcgc aatttttag

Secuencia de aminoácidos de ca A/ck/Hong Kong/G9/97 N2 (SEC ID N° 20) Longitud de la molécula completa: 469 aa

1 mnpnqkii ai gsvltiati cllmqiaila ttmtlhfkqn ecinssnnqv  
 51 vpcepiier niteivhl ns ttlekeicpk vadyrnwskp qcqitgfapf  
 101 skdnsirlsa ggdiwvtr ep yvscglgkcy qfalgggttl enkhsngtah  
 151 drtphrllm nelgvpfhla tkqvciawss sschdgkaw l hvcvtgddrn  
 201 atasiiydgi lvdsigswsk nilrtqesec vcingtcavv mtdgsasgra  
 251 dtrilfireg kiahisplsg saqhveecsc yprypevrcv crdnwksnr  
 301 pvlyinmany sidssyvcsg lvgdtprrnd rssssncrdp nnergapgvk  
 351 gwafdngndi wmgtrikkds rsgyetfrvi ggwtansks qinrqvlds  
 401 dnssgysgif svegkscinr cfyvelirgr pketrvwts nsiivfcgts  
 451 gtygtgswpd ganinfmpi

Resumen de designaciones de sec id n°

| SEC ID N°    | HA o NA | NOMBRE DE LA CEPA              | Aminoácidos o nucleótidos |
|--------------|---------|--------------------------------|---------------------------|
| SEC ID N° 1  | HA (H5) | ca A/Vietnam/1203/04           | Nucleótidos               |
| SEC ID N° 2  | NA (N1) | ca A/Vietnam/1203/04           | Nucleótidos               |
| SEC ID N° 3  | HA (H5) | ca A/Hong Kong/213/03          | Nucleótidos               |
| SEC ID N° 4  | NA (N1) | ca A/Hong Kong/213/03          | Nucleótidos               |
| SEC ID N° 5  | HA (H5) | ca A/Hong Kong/491/97          | Nucleótidos               |
| SEC ID N° 6  | NA (N1) | ca A/Hong Kong/486/97          | Nucleótidos               |
| SEC ID N° 7  | HA (H5) | ca A/Hong Kong/491/97 (Ser211) | Nucleótidos               |
| SEC ID N° 8  | NA (N1) | ca A/Hong Kong/486/97          | Nucleótidos               |
| SEC ID N° 9  | HA (H9) | ca A/ck/Hong Kong/G9/97        | Nucleótidos               |
| SEC ID N° 10 | NA (N2) | ca A/ck/Hong Kong/G9/97        | Nucleótidos               |
| SEC ID N° 11 | HA (H5) | ca A/Vietnam/1203/04           | Aminoácidos               |
| SEC ID N° 12 | NA (N1) | ca A/Vietnam/1203/04           | Aminoácidos               |
| SEC ID N° 13 | HA (H5) | ca A/Hong Kong/213/03          | Aminoácidos               |
| SEC ID N° 14 | NA (N1) | ca A/Hong Kong/213/03          | Aminoácidos               |
| SEC ID N° 15 | HA (H5) | ca A/Hong Kong/491/97          | Aminoácidos               |
| SEC ID N° 16 | NA (N1) | ca A/Hong Kong/486/97          | Aminoácidos               |
| SEC ID N° 17 | HA (H5) | ca A/Hong Kong/491/97 (Ser211) | Aminoácidos               |

|              |         |                         |             |
|--------------|---------|-------------------------|-------------|
| SEC ID Nº 18 | NA (N1) | ca A/Hong Kong/486/97   | Aminoácidos |
| SEC ID Nº 19 | HA (H9) | ca A/ck/Hong Kong/G9/97 | Aminoácidos |
| SEC ID Nº 20 | NA (N2) | ca A/ck/Hong Kong/G9/97 | Aminoácidos |

**Listado de secuencias**

5 <110> MedImmune Vaccines, Inc et al.  
 <120> VARIANTES DE LA HEMAGLUTININA Y LA NEURAMIDINASA DE INFLUENZA  
 <130> S2570 EP/1  
 <150> 60/574,553  
 <151> 2004-05-25  
 <150> 60/657,554  
 10 <151> 2005-02-28  
 <160> 20  
 <170> PatentIn version 3.3  
 <210> 1  
 <211> 1767  
 15 <212> ADN  
 <213> Virus Influenza  
 <400> 1  
 agcaaaagca ggggttcaat ctgtcaaaat ggagaaaata gtgcttcttt ttgcaatagt 60  
 cagtcttgtt aaaagtgatc agatttgcat tggttacat gcaaacaact cgacagagca 120  
 ggttgacaca ataatggaaa agaacgttac tgttacacat gccaagaca tactggaaaa 180  
 gaaacacaac gggaagctct gcgatctaga tggagtgaag cctctaattt tgagagattg 240  
 tagcgtagct ggatggctcc tcggaaaccc aatgtgtgac gaattcatca atgtgccgga 300  
 atggtcttac atagtggaga aggccaatcc agtcaatgac ctctgttacc caggggattt 360  
 caatgactat gaagaattga aacacctatt gagcagaata aaccattttg agaaaattca 420  
 gatcatcccc aaaagtctct ggtccagtca tgaagcctca ttaggggtga gctcagcatg 480  
 tccataccag ggaaagtcct cttttttcag aaatgtggtg tggcttatca aaaagaacag 540  
 tacataccca acaataaaga ggagctacaa taataccaac caagaagatc ttttgggtact 600  
 gtgggggatt caccatccta atgatgcggc agagcagaca aagctctatc aaaaccaaac 660  
 cacctatatt tccgttggga catcaacact aaaccagaga ttggtacca gaatagctac 720  
 tagatccaaa gtaaacgggc aaagtgaag gatggagttc ttctggacaa ttttaagcc 780  
 gaatgatgca atcaacttcg agagtaatgg aaatttcatt gctccagaat atgcatacaa 840  
 aattgtcaag aaaggggact caacaattat gaaaagtga ttggaatatg gtaactgcaa 900  
 caccaagtgt caaactccaa tgggggcgat aaactctagc atgccattcc acaatatata 960  
 ccctctcacc attggggaat gcccacaata tgtgaaatca aacagattag tccttgccgac 1020  
 tgggctcaga aatagccctc aaagagagac tcgaggatta tttggagcta tagcaggttt 1080  
 tatagagga ggatggcagg gaatggtaga tggttggtat gggtagacc atagcaatga 1140  
 gcaggggagt ggttacgctg cagacaaaga atccactcaa aaggcaatag atggagtcac 1200  
 caataaggct aactcgatca ttgacaaaat gaacactcag tttgaggccg ttggaagga 1260

atttaacaac ttagaaagga gaatagagaa tttaaacaag aagatggaag acgggttcct 1320  
 agatgtctgg acttataatg ctgaacttct ggttctcatg gaaaatgaga gaactctaga 1380  
 ctttcatgac tcaaatgtca agaaccctta cgacaaggtc cgactacagc ttagggataa 1440  
 tgcaaaggag ctgggtaacg gttgtttcga gttctatcat aaatgtgata atgaatgtat 1500  
 ggaaagtgtg agaaatggaa cgtatgacta cccgcagtat tcagaagaag cgagactaaa 1560  
 aagagaggaa ataagtggag taaaattgga atcaatagga atttaccaaa tactgtcaat 1620  
 ttattctaca gtggcgagtt ccctagcact ggcaatcatg gtagctggtc tatccttatg 1680  
 gatgtgctcc aatgggtcgt tacaatgcag aatttgcatt taaatttgtg agttcagatt 1740  
 gtagttaaaa acacccttgt ttctact 1767

<210> 2

<211> 1398

<212> ADN

5 <213> Virus Influenza

<400> 2

agcaaaaagca ggagttcaaa atgaatccaa atcagaagat aataaccatc ggggtcaatct 60  
 gtatggtaac tggaatagtt agcttaatgt taaaaattgg gaacatgatc tcaatatggg 120  
 tcagtcattc aattcacaca gggaatcaac accaatctga accaatcagc aataactaatt 180  
 ttcttactga gaaagctgtg gcttcagtaa aattagcggg caattcatct ctttgcccca 240  
 ttaacggatg ggctgtatac agtaaggaca acagtataag gatcggttcc aagggggatg 300  
 tgtttgttat aagagagccg ttcattctcat gctcccactt ggaatgcaga actttctttt 360  
 tgactcaggg agccttgctg aatgacaagc actccaatgg gactgtcaaa gacagaagcc 420  
 ctcacagaac attaatgagt tgtcctgtgg gtgaggctcc ctccccatat aactcaaggt 480  
 ttgagtctgt tgcttgggtca gcaagtgctt gccatgatgg caccagttgg ttgacgattg 540  
 gaatttctgg ccagacaat ggggctgtgg ctgtattgaa atacaatggc ataataacag 600  
 acactatcaa gagttggagg aacaacatac tgagaactca agagtctgaa tgtgcatgtg 660  
 taaatggctc ttgctttact gtaatgactg acggaccaag taatggtcag gcatcacata 720  
 agatcttcaa aatggaaaaa gggaaagtgg ttaaatcagt cgaattggat gctcctaatt 780  
 atcactatga ggaatgctcc tgttatccta atgccggaga aatcacatgt gtgtgcaggg 840  
 ataattggca tggtccaaat cggccatggg tatctttcaa tcaaaatttg gagtatcaaa 900  
 taggatatat atgcagtgga gttttcggag acaatccacg cccaatgat ggaacaggta 960  
 gttgtggtcc ggtgtcctct aacggggcat atggggtaaa agggttttca tttaaatacg 1020  
 gcaatggtgt ctggatcggg agaaccaaaa gcactaattc caggagcggc tttgaaatga 1080  
 tttgggatcc aaatgggtgg actgaaacgg acagtagctt ttcagtgaag caagatatcg 1140  
 tagcaataac tgattgggtca ggatatagcg ggagttttgt ccagcatcca gaactgacag 1200  
 gactagattg cataagacct tgtttctggg ttgagttgat cagagggcgg cccaaagaga 1260  
 gcacaatttg gactagtggg agcagcatat ctttttgtgg tgtaaatagt gacactgtgg 1320  
 gttggtcttg gccagacggt gctgagttgc cattcaccat tgacaagtag tttgttcaaa 1380  
 aaactccttg ttctact 1398

<210> 3

<211> 1767

<212> ADNA

10

&lt;213&gt; Virus Influenza

&lt;400&gt; 3

|                                                                   |      |
|-------------------------------------------------------------------|------|
| agcaaaagca ggggttcaat ctgtcaaaat ggagaaaata gtgcttcttt ttgcaatagt | 60   |
| cagtcttggt aaaagtgatc agatttgcat tggttaccat gcaaacaact cgacagagca | 120  |
| ggttgacaca ataatggaaa agaacgttac tgttacacat gccaagaca tactggaaaa  | 180  |
| gacacacaac gggaagctct gcgatctaga tggagtgaag cctctaattt tgagagattg | 240  |
| tagtgtagct ggatggctcc tcggaaaccc aatgtgtgac gaattcatca atgtgccgga | 300  |
| atggtcttac atagtggaga aggccaatcc agccaatgac ctctgttacc caggggattt | 360  |
| caacgactat gaagaattga aacacctatt gagcagaata aaccattttg agaaaattca | 420  |
| gatcatcccc aaaaattctt ggtccagtc tgaagcctca ttaggggtga gctcagcatg  | 480  |
| tccataccaa ggaaagtctt cctttttcag gaatgtggt tggcttatca aaaagaacaa  | 540  |
| tgcataccca acaataaaga ggagctacaa taataccaac caagaagatc ttttggtatt | 600  |
| gtgggggatt caccatccta atgatgcggc agagcagact aggtctatc aaaaccaac   | 660  |
| cacctacatt tccgttggga catcaacact aaaccagaga ttggtaccaa aaatagctac | 720  |
| tagatccaaa gtaaacgggc aaaatggaag gatggagtcc ttctggacaa ttttaaaacc | 780  |
| gaatgatgca atcaacttcg agagcaatgg aaatttcatt gctccagaat atgcatacaa | 840  |
| aattgtcaag aaaggggact cagcaattat gaaaagtga ttggaatatg gtaactgcaa  | 900  |
| caccaagtgt caaactccaa tgggggcat aaactctagt atgccattcc acaatataca  | 960  |
| ccctctcacc atcggggaat gcccacaata tgtgaaatca aacagattag tccttgcgac | 1020 |
| tgggctcaga aatagccctc aaagagagac tcgaggatta tttggagcta tagcaggttt | 1080 |
| tatagagggg ggatggcagg gaatggtaga tggttggtat gggtaccacc atagcaatga | 1140 |
| gcaggggagt gggtagctg cagacaaaga atccactcaa aaggcaatag atggagtac   | 1200 |
| caataaggtc aactcgatca ttgacaaaat gaacactcag tttgaggccg ttggaaggga | 1260 |
| atttaataac ttgaaagga gaatagagaa tttaacaag aagatggaag acggattcct   | 1320 |
| agatgtctgg acttataatg ctgaacttct ggttctcatg gaaaatgaga gaactctaga | 1380 |
| ctttcatgac tcaaatgtca agaaccttta cgacaaggtc cgactacagc ttagggataa | 1440 |
| tgcaaaggag ctgggtaacg gttgtttcga gttctatcac aaatgtgata atgaatgtat | 1500 |
| ggaaagtgtg agaaacggaa cgtatgacta cccgcagtat tcagaagaag caagactaaa | 1560 |
| aagagaggaa ataagtggag taaaattgga gtcaatagga acttaccaaa tactgtcaat | 1620 |
| ttattctaca gtggcgagtt ccctagcact ggcaatcatg gtagctggtc tatctttatg | 1680 |
| gatgtgctcc aatgggtcgt tacaatgcag aatttgcatt taaatttgtg agttcagatt | 1740 |
| gtagttaaaa acacccttgt ttctact                                     | 1767 |

5

&lt;210&gt; 4

&lt;211&gt; 1458

&lt;212&gt; ADN

&lt;213&gt; Virus Influenza

&lt;400&gt; 4

```

agcaaaagca ggagttcaaa atgaatccaa atcagaagat aacaaccatt ggatcaatct      60
gtatggtaat tggaatagtt agcttgatgt tacaaattgg gaacataatc tcaatatggg      120
ttagtcattc aattcaaaca gggaatcaac accaggctga accatgcaat caaagcatta      180
ttacttatga aaacaacacc tgggtaaacc agacatatgt caacatcagc aataccaatt      240
ttcttactga gaaagctgtg gcttcagtaa cattagcggg caattcatct ctttgcccca      300
ttagtggtat ggctgtatac agtaaggaca acggtataag aatcgggtcc aagggggatg      360
tgtttggtat aagagagccg ttcatctcat gctcccaactt ggaatgcaga actttctttt      420
tgactcaggg agccttgctg aatgacaagc attctaattg gaccgtcaaa gacagaagcc      480
ctcacagaac attaatgagt tgtcccgtag gtgaggtcc ttccccatac aactcgaggt      540
ttgagtcgtg tgcttggtcg gcaagtgcct gtcagtatgg cactagttag ttgacaattg      600
gaatttctgg ccagacaat ggggctgtgg ctgtattgaa atacaatggc ataataacag      660
acactatcaa gagtggagg aacaacataa tgagaactca agagtctgaa tgtgcatgtg      720
taaattggctc ttgctttact gttatgactg atggaccaag taatgggcag gcttcataca      780
aaatcttcag aatagaaaaa gggaaagtag ttaaatacgc cgaattaaat gccctaatt      840
atcactatga ggagtgctcc tggtatcctg atgctggaga aatcacatgt gtgtgcaggg      900
ataactggca tggctcaaat cggccatggg tatctttcaa tcaaaatttg gagtatcgaa      960
taggatatat atgcagtgga gttttcggag acaatccacg cccaatgat gggacaggca     1020
gttgtgtgcc ggtgtccct aaaggggcat atggaataaa agggttctca tttaaatacg     1080
gcaatggtgt ttggatcggg agaaccaaaa gcactaattc caggagcggc tttgaaatga     1140
tttgggatcc aaatggatgg actggtacgg acagtaattt ttcagtaaag caagatattg     1200
tagctataac cgattggtca ggatatagcg ggagttttgt ccagcatcca gaactgacag     1260
gattagattg cataagacct tgtttctggg ttgagctaag cagagggcgg cccaagaga     1320
gcacaatttg gactagtggg agcagcatat cttttgtgg tgtaaatagt gacactgtgg     1380
gttggtcttg gccagacggt gctgagttgc cattcaccat tgacaagtag tttgttcaaa     1440
aaactccttg tttctact

```

&lt;210&gt; 5

&lt;211&gt; 1767

5 &lt;212&gt; ADN

&lt;213&gt; Virus influenza

&lt;400&gt; 5

|                                                                     |      |
|---------------------------------------------------------------------|------|
| agcaaaagca ggggtataat ctgtcaaaat ggagaaaata gtgcttcttc ttgcaacagt   | 60   |
| cagccttggtt aaaagtgacc agatttgcat tggttacccat gcaaacaact cgacagagca | 120  |
| agttgacaca ataatggaaa agaattgttac tgttacacat gcccaagaca tactggaaaag | 180  |
| gacacacaac ggggaagctct gcgatctaaa tggagtgaag cctctgattt tgagggattg  | 240  |
| tagtgtagct ggatggctcc tcggaaaccc tatgtgtgac gaattcatca atgtgccgga   | 300  |
| atggtcttac atagtggaga aggccagtc agccaatgac ctctgttatc cagggaattt    | 360  |
| caacgactat gaagaactga aacacctatt gagcagaata aaccattttg agaaaattca   | 420  |
| gataatcccc aaaagttctt ggtccaatca tgatgcctca tcaggggtga gctcagcatg   | 480  |
| tccatacctt gggaggtcct cctttttcag aaatgtggtt tggcttatca aaaagaacag   | 540  |
| tagctaccca acaataaaga ggagctacaa taataccaac caagaagatc ttttgggtact  | 600  |
| gtgggggatt caccatccta atgatgcggc agagcagaca aggtctatc aaaacccaac    | 660  |
| cacctacatt tccgttgga catcaacact gaaccagaga ttggttccag aaatagctac    | 720  |
| tagacccaaa gtaaacgggc aaagtggaa aatggagttc ttctggacaa ttttaagcc     | 780  |
| gaatgatgcc atcaatttcg agagtaatgg aaatttcatt gctccagaat atgcatacaa   | 840  |
| aattgtcaag aaaggggact caacaattat gaaaagtga ttggaatatg gtaactgcaa    | 900  |
| caccaagtgt caaactccaa tgggggcaat aaactctagt atgccattcc acaacataca   | 960  |
| ccccctcacc atcggggaat gcccacaata tgtgaaatca aacagattag tccttgcaac   | 1020 |
| tggactcaga aataccctc aacgagagac gcgaggacta tttggagcta tagcaggttt    | 1080 |
| tatagaggga ggatggcagg gaatggtaga tggttggtat ggggtaccacc atagcaatga  | 1140 |
| gcaggggagt ggatacgtc cagaccaaga atccacacaa aaggcaatag atggagtcac    | 1200 |
| caataaggtc aactcgatca ttaacaaaat gaacactcag tttgaggccg ttggaaggga   | 1260 |
| atttaataac ttggaaagga ggatagagaa tttaacaag aaaatggaag acggattcct    | 1320 |
| agatgtctgg acttacaatg ccgaacttct ggttctcatg gaaaatgaga gaactctaga   | 1380 |
| ctttcatgac tcaaatgtca agaaccttta cgacaaggtc cgactacagc ttagggataa   | 1440 |
| tgcaaaggag ctgggtaatg gttgtttcga attctatcac aaatgtgata acgaatgtat   | 1500 |
| ggaaagtgt aaaaacggaa cgtatgacta cccgcagtat tcagaagaag caagactaaa    | 1560 |
| cagagaggaa ataagtggag taaaattgga atcaatggga acttaccaa tactgtcaat    | 1620 |
| ttattcaaca gtggcgagtt ccctagcact ggcaatcatg gtagctggtc tatctttatg   | 1680 |
| gatgtgctcc aatggatcgt tacaatgcag aatttgcat taaatttggtg agttcagatt   | 1740 |
| gtagttaaaa acacccttgt ttctact                                       | 1767 |

&lt;210&gt; 6

&lt;211&gt; 1401

&lt;212&gt; ADN

&lt;213&gt; Virus influenza

&lt;400&gt; 6



```

agcaaaagca ggagttttaa atgaatccaa atcagaagat aataaccatt ggatcaatct      60
gcatggtagt tgggataatc agcttgatgt tacaaattgg aaacacaata tcagtatggg      120
tcagccacat aattaaaact tggcacccaa accagcctga accatgcaac caaagcatca      180
atttttacac tgagcaggct gcagcttcag tgacattagc gggcaattcc tctctctgcc      240
ctattagtgg atgggctata tacagcaagg acaatagtat aagaattggg tccaaagggg      300
atgtgtttgt tataagagaa ccattcatct catgctccca tttggaatgc agaacccttt      360
tcttgaccca aggagcccta ttgaatgaca agcattctaa tgggaccgtc aaagacagga      420
gcccctatag aactttaatg agctgtcctg ttgggtaggc cccttcccca tacaactcaa      480
ggtttgagtc tgttgcttgg tcagcaagtg cttgccatga tggcattagt tggctaacaa      540
ttggaatttc cggtcgggat aatggggctg tggctgtgtt gaaatacaat ggcataataa      600
cagacaccat caagagttag aggaacaaca cactgaggac gcaagagtct gaatgtgcat      660
gtgtgaatgg ttcttgtttt actgtaatga cagatggacc gagtaatgaa caggcctcat      720
acaagatttt caagatagaa aaggggaggg tagtcaaatc agttgagttg aacgccccta      780
attatcatta cgaggaatgc tcctgttatc ctgatgctgg cgaaatcaca tgtgtgtgca      840
gggataattg gcatggctcg aaccgaccat ggggtgtctt caatcagaat ctggagtatc      900
aaataggata tatatgcagt ggggttttct gagacagtcc acgcccctaat gatgggacag      960
gcagtgtggg tccagtgtct cttaacggag cgtatggagt aaaagggttt tcatttaaat     1020
acggcaatgg tgtttggatc gggagaacca aaagcactag ttccaggagc ggttttgaaa     1080
tgatttgagg tccaaatggg tggaccgaaa cagacagtag cttctcgttg aagcaagaca     1140
tcatagcgat aactgattgg tcaggataca gcgggagttt tattcaacat ccagaactga     1200
caggattaaa ttgcatgaga ccttgcttct ggggtgaact aatcagaggg aggcccaaag     1260
agaaaacaat ctggactagt gggagcagta tatctttctg tgggtgtaaat agtgacactg     1320
tgggttggtc ttggccagac ggtgctgagt tgccatacac cattgacaag tagtttgttc     1380
aaaaaactcc ttgtttctac t                                             1401

```

&lt;210&gt; 7

&lt;211&gt; 1767

&lt;212&gt; ADN

5 &lt;213&gt; Virus influenza

&lt;400&gt; 7

```

agcaaaagca ggggtataat ctgtcaaaat ggagaaaata gtgcttcttc ttgcaacagt      60
cagccttggt aaaagtgacc agatttgcat tggttaccat gcaaacaact cgacagagca      120
agttgacaca ataatggaaa agaatgttac tgttacacat gcccaagaca tactggaaaag      180
gacacacaac gggaagctct cgcgtctaaa tggagtgaag cctctgattt tgagggattg      240
tagtgtagct ggatgggtcc tcggaaaccc tatgtgtgac gaattcatca atgtgccgga      300
atggtcttac atagtggaga aggccagtcc agccaatgac ctctgttatc cagggaattt      360
caacgactat gaagaactga aacacctatt gagcagaata aaccattttg agaaaattca      420
gataatcccc aaaagtctct ggtccaatca tgatgcctca tcaggggtga gctcagcatg      480
tccatacctt gggaggtcct cctttttcag aaatgtggtg tggcttatca aaaagaacag      540
tagctacca acaataaaga ggagctacaa taataccaac caagaagatc ttttggtagt      600
gtgggggatt caccatccta atgatgcggc agagcagaca aggtctctac aaaacccaac      660
cacctacatt tccgttgga catcaacact gaaccagaga ttggtttcag aaatagctac      720
tagacccaaa gtaaacgggc aaagtggaa atgtggagtc ttctggacaa ttttaagcc      780
gaatgatgcc atcaatttcg agagtaatgg aaatttcatt gctccagaat atgcatacaa      840
aattgtcaag aaaggggact caacaattat gaaaagtga ttggaatatg gtaactgcaa      900
caccaagtgt caaactccaa tgggggcaat aaactctagt atgccattcc acaacataca      960
ccccctcacc atcggggaat gcccacaata tgtgaaatca aacagattag tccttgcaac     1020
tggactcaga aatacccctc aacgagagac gcgaggacta tttggagcta tagcaggttt     1080
tatagaggga ggatggcagg gaatggtaga tggttggtat gggtagacc atagcaatga     1140
gcaggggagt ggatacgtg cagaccaaga atccacacaa aaggcaatag atggagtcac     1200
caataagggt aactcgatca ttaacaaaat gaacactcag tttgaggccg ttggaaggga     1260
atttaataac ttggaaagga ggatagagaa tttaaacaag aaaatggaag acggattcct     1320
agatgtctgg acttacaatg ccgaacttct ggttctcatg gaaaatgaga gaactctaga     1380
ctttcatgac tcaaatgtca agaaccttta cgacaaggtc cgactacagc ttagggataa     1440
tgcaaaggag ctgggtaatg gttgtttcga attctatcac aaatgtgata acgaatgtat     1500
ggaaagtgtg aaaaacggaa cgtatgacta cccgcagtat tcagaagaag caagactaaa     1560
cagagaggga ataagtggag taaaattgga atcaatggga acttaccaa tactgtcaat     1620
ttattcaaca gtggcgagtt ccctagcact ggcaatcatg gtagctggtc tatctttatg     1680
gatgtgctcc aatggatcgt tacaatgcag aatttgcatt taaatttgtg agttcagatt     1740
gtagttaaaa acacccttgt ttctact                                         1767
<210> 8
<211> 1401
<212> ADN
5 <213> Virus influenza
<400> 8

```

```

agcaaaagca ggagttttaa atgaatccaa atcagaagat aataaccatt ggatcaatct    60
gcatggtagt tgggataatc agcttgatgt tacaaattgg aaacacaata tcagtatggg    120
tcagccacat aattaaaact tggcacccaa accagcctga accatgcaac caaagcatca    180
atTTTTacac tgagcaggct gcagcttcag tgacattagc gggcaattcc tctctctgcc    240
ctattagtgg atgggtata tacagcaagg acaatagtat aagaattggg tccaaagggg    300
atgtgtttgt tataagagaa ccattcatct catgctccca tttggaatgc agaacctttt    360
tcttgaccca aggagcccta ttgaatgaca agcattctaa tgggaccgtc aaagacagga    420
gcccctatag aactttaatg agctgtcctg ttgggtaggc cccttcccca tacaactcaa    480
ggtttgagtc tgttgcttgg tcagcaagtg cttgccatga tggcattagt tggctaacaa    540
ttggaatttc cggtcoggat aatggggctg tggctgtgtt gaaatacaat ggcataataa    600
cagacaccat caagagtgg aggaacaaca cactgaggac gcaagagtct gaatgtgcat    660
gtgtgaatgg ttcttgtttt actgtaatga cagatggacc gagtaatgaa caggcctcat    720
acaagatttt caagatagaa aaggggaggg tagtcaaatc agttgagttg aacgccccta    780
attatcatta cgaggaatgc tcctgttatc ctgatgctgg cgaaatcaca tgtgtgtgca    840
gggataattg gcatggctcg aaccgaccat ggggtgtctt caatcagaat ctggagtatc    900
aaataggata tatatgcagt ggggttttcg gagacagtcc acgcccgaat gatgggacag    960
gcagttgtgg tccagtgtct cttaacggag cgtatggagt aaaagggttt tcatttaaat 1020
acggcaatgg tgtttggatc gggagaacca aaagcactag ttccaggagc ggttttgaaa 1080
tgatttggga tccaaatggg tggaccgaaa cagacagtag cttctcgttg aagcaagaca 1140
tcatagcgat aactgattgg tcaggatata gcgggagttt tattcaacat ccagaactga 1200
caggattaaa ttgcatgaga ccttgcttct gggttgaact aatcagaggg aggcccaaag 1260
agaaaacaat ctggactagt gggagcagta tatctttctg tgggtgtaaat agtgacactg 1320
tgggttggtc ttggccagac ggtgctgagt tgccatacac cattgacaag tagtttgttc 1380
aaaaaactcc ttgtttctac t                                         1401

```

&lt;210&gt; 9

&lt;211&gt; 1690

&lt;212&gt; ADN

5 &lt;213&gt; Virus influenza

&lt;400&gt; 9

```

ttaaccactc aagatggaag caataccact aataactata ctactagtag taacagcaag    60
caatgcagac aaaatctgca tcggctacca atcaacaaac tccacagaaa ccgtagacac    120
gctaacagaa aacaatgttc ctgtgacaca tgccaaagaa ttgtccaca cagagcacia    180

```

|                                                                    |      |
|--------------------------------------------------------------------|------|
| tgggatgctg tgtgcaacaa atctgggacg tcctcttatt ctagacactt gcaccattga  | 240  |
| aggactgac tatggcaacc cttcttgtga tctactgttg ggaggaagag aatggtccta   | 300  |
| catcgtcgaa agaccatcgg ctgttaatgg aatgtgttac cccgggaatg tagaaaacct  | 360  |
| agaggaacta aggtcatttt ttagttctgc tagttcctac caaagaatcc agatctttcc  | 420  |
| agacacaatc tggaatgtgt cttacagtgg aacaagcaaa gcatgttcag attcattcta  | 480  |
| caggagcatg agatgggtga ctcaaaagaa caacgcttac cctattcaag acgccaata   | 540  |
| cacaaataat agaggaaaga gcattctttt catgtggggc ataatcacc cacctaccga   | 600  |
| tactgcacag acaaatctgt acacaaggac tgacacaaca acaagtgtgg caacagaaga  | 660  |
| tataaatagg accttcaaac cagtgatagg gccaaagccc cttgtcaatg gtctgcaggg  | 720  |
| aagaattgat tattattggt cggattgaa accaggtcag acattgcgag taagatccaa   | 780  |
| tgggaatcta atcgctccat ggtatgggca cattctttca ggagagagcc acggaagaat  | 840  |
| cctgaagact gatttaaaca gtggtagctg tgtagtcaa tgtcaaacag aaagaggtgg   | 900  |
| cttaataact actttgccat tccacaatgt cagtaaataat gcatttgga actgccaaa   | 960  |
| atatgttga gtaaagagtc tcaaaactggc agttggtctg aggaatgtgc ctgctagatc  | 1020 |
| aagtagagga ctatttgggg ccatagctgg attcatagag ggaggttgg cagggctggt   | 1080 |
| cgctggttg tatgggttcc agcattcaaa tgatcaaggg gtggtatag ctgcagatag    | 1140 |
| agactcaact caaagggcaa ttgacaaaat aacgtccaaa gtgaataata tagtcgataa  | 1200 |
| aatgaacaag caatatgaaa ttattgatca tgaattcagc gaggttgaaa atagactcaa  | 1260 |
| tatgatcaat aataagattg atgaccagat acaagacata tgggcatata acgctgaatt  | 1320 |
| gctagtgtctg cttgaaaacc agaaaacact cgatgagcat gatgcgaatg taaacaatct | 1380 |
| atataacaaa gtgaagaggg cactgggttc caatgcaatg gaagatggga aaggatgttt  | 1440 |
| cgagctatac cataaatgtg atgatcagt catggagaca attcggaacg ggacctataa   | 1500 |
| caggaggaag tataaagagg aatcaagact agaaagacag aaaatagaag gggtaagct   | 1560 |
| ggaatctgaa ggaacttaca aaatcctcac catttattcg actgtgcct catctcttgt   | 1620 |
| gattgcaatg gggtttgctg ccttcttgtt ctgggccatg tccaatggat cttgcagatg  | 1680 |
| caacatttga                                                         | 1690 |
| <210> 10                                                           |      |
| <211> 1428                                                         |      |
| <212> ADN                                                          |      |
| 5 <213> Virus influenza                                            |      |
| <400> 10                                                           |      |
| aaatgaatcc aaatcagaag ataatagcaa ttggctctgt ttctctaact attgcgacaa  | 60   |
| tatgcctcct catgcagatt gctatcttag caacgactat gacactacat ttcaagcaga  | 120  |
| atgaatgcat caactcctcg aataatcaag tagtgccatg tgaaccaatc ataatagaaa  | 180  |

```

ggaacataac agagatagtg catttgaata gtactacctt agagaaggaa atttgtccta 240
aagtagcaga ctacaggaat tggcctaaaac cacaatgtca aatcacaggg ttogctcctt 300
tctccaagga caattcaatt aggcctctccg cagggtggaga tatttgggtg acaagagAAC 360
cttatgtatc gtgcggtctt ggtaaagtgt atcaatttgc acttgggcag ggaaccactt 420
tggagaacaa acaactcaaac ggcacagcac atgatagaac tcctcataga acccttttaa 480
tgaatgagtt ggggtgttccg ttccatttgg caaccAAACA agtgtgcata gcatggtcca 540
gctcaagctg ccatgatggg aaagcatggt tacatgtttg tgtcactggg gatgatagaa 600
atgcaacggc tagcatcatt tatgatggga tacttgttga cagtattggt tcatggtcta 660
aaaacatcct cagaactcag gagtcagaat gcgtttgcat caatggaacc tgtgcagtag 720
taatgactga tggaaagtga tcaggaaggg ctgacactag aatactattt attagagagg 780
ggaaaattgc acacattagc ccattgtcag gaagtgtcga gcatgtggag gaatgtcct 840
gttaccctcg atatccagaa gttagatgtg tttgcagaga caattggaag ggatccaata 900
ggcccggtct atatataaat atggcAAATT atagtattga ttccagttat gtgtgctcag 960
gacttgttgg cgacacacca agaaatgatg ataggtctag cagcagcaac tgcagagatc 1020
ctaataacga gagagggggc ccaggagtaa aagggtgggc ctttgacaat ggaaatgaca 1080
tttgatggg aagaacaatc aaaaaggatt cgcgctcagg ttatgagact ttcagggtca 1140
ttggtggttg gaccactgct aattccaagt cacagataaa tagacaagtc atagttgaca 1200
gtgataactc gtctgggtat tctggtatct tctctgttga aggcaaaagc tgcataca 1260
ggtgttttta cgtggagttg ataagaggaa gaccaaagga gactaggggtg tgggtgactt 1320
caaatagcat cattgtattt tgtggaactt caggtaacct tggAACAGGC tcatggcctg 1380
atggggcgaa tatcaatttc atgcctatat aagctttcgc aatttttag 1428

```

&lt;210&gt; 11

&lt;211&gt; 564

&lt;212&gt; PRT

5 &lt;213&gt; Virus influenza

&lt;400&gt; 11

```

Met Glu Lys Ile Val Leu Leu Phe Ala Ile Val Ser Leu Val Lys Ser
1           5           10          15

```

```

Asp Gln Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Glu Gln Val
20           25           30

```

```

Asp Thr Ile Met Glu Lys Asn Val Thr Val Thr His Ala Gln Asp Ile
35           40           45

```

```

Leu Glu Lys Lys His Asn Gly Lys Leu Cys Asp Leu Asp Gly Val Lys
50           55           60

```

Pro Leu Ile Leu Arg Asp Cys Ser Val Ala Gly Trp Leu Leu Gly Asn  
 65 70 75 80

Pro Met Cys Asp Glu Phe Ile Asn Val Pro Glu Trp Ser Tyr Ile Val  
 85 90 95

Glu Lys Ala Asn Pro Val Asn Asp Leu Cys Tyr Pro Gly Asp Phe Asn  
 100 105 110

Asp Tyr Glu Glu Leu Lys His Leu Leu Ser Arg Ile Asn His Phe Glu  
 115 120 125

Lys Ile Gln Ile Ile Pro Lys Ser Ser Trp Ser Ser His Glu Ala Ser  
 130 135 140

Leu Gly Val Ser Ser Ala Cys Pro Tyr Gln Gly Lys Ser Ser Phe Phe  
 145 150 155 160

Arg Asn Val Val Trp Leu Ile Lys Lys Asn Ser Thr Tyr Pro Thr Ile  
 165 170 175

Lys Arg Ser Tyr Asn Asn Thr Asn Gln Glu Asp Leu Leu Val Leu Trp  
 180 185 190

Gly Ile His His Pro Asn Asp Ala Ala Glu Gln Thr Lys Leu Tyr Gln  
 195 200 205

Asn Pro Thr Thr Tyr Ile Ser Val Gly Thr Ser Thr Leu Asn Gln Arg  
 210 215 220

Leu Val Pro Arg Ile Ala Thr Arg Ser Lys Val Asn Gly Gln Ser Gly  
 225 230 235 240

Arg Met Glu Phe Phe Trp Thr Ile Leu Lys Pro Asn Asp Ala Ile Asn  
 245 250 255

Phe Glu Ser Asn Gly Asn Phe Ile Ala Pro Glu Tyr Ala Tyr Lys Ile  
 260 265 270

Val Lys Lys Gly Asp Ser Thr Ile Met Lys Ser Glu Leu Glu Tyr Gly  
 275 280 285

Asn Cys Asn Thr Lys Cys Gln Thr Pro Met Gly Ala Ile Asn Ser Ser  
 290 295 300

Met Pro Phe His Asn Ile His Pro Leu Thr Ile Gly Glu Cys Pro Lys  
 305 310 315 320

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

Pro Gln Arg Glu Thr Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile  
340 345 350

Glu Gly Gly Trp Gln Gly Met Val Asp Gly Trp Tyr Gly Tyr His His  
355 360 365

Ser Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Lys Glu Ser Thr Gln  
370 375 380

Lys Ala Ile Asp Gly Val Thr Asn Lys Val Asn Ser Ile Ile Asp Lys  
385 390 395 400

Met Asn Thr Gln Phe Glu Ala Val Gly Arg Glu Phe Asn Asn Leu Glu  
405 410 415

Arg Arg Ile Glu Asn Leu Asn Lys Lys Met Glu Asp Gly Phe Leu Asp  
420 425 430

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Val | Trp | Thr | Tyr | Asn | Ala | Glu | Leu | Leu | Val | Leu | Met | Glu | Asn | Glu | Arg |
|     |     | 435 |     |     |     |     | 440 |     |     |     |     | 445 |     |     |     |

Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu Tyr Asp Lys Val  
450 455 460

Arg Leu Gln Leu Arg Asp Asn Ala Lys Glu Leu Gly Asn Gly Cys Phe  
465 470 475 480

Glu Phe Tyr His Lys Cys Asp Asn Glu Cys Met Glu Ser Val Arg Asn  
485 490 495

Gly Thr Tyr Asp Tyr Pro Gln Tyr Ser Glu Glu Ala Arg Leu Lys Arg  
500 505 510

Glu Glu Ile Ser Gly Val Lys Leu Glu Ser Ile Gly Ile Tyr Gln Ile  
515 520 525

Leu Ser Ile Tyr Ser Thr Val Ala Ser Ser Leu Ala Leu Ala Ile Met  
530 535 540

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Val | Ala | Gly | Leu | Ser | Leu | Trp | Met | Cys | Ser | Asn | Gly | Ser | Leu | Gln | Cys |
| 545 |     |     |     |     | 550 |     |     |     |     | 555 |     |     |     |     | 560 |

Arg Ile Cys Ile

$\langle 210 \rangle$  12

<211> 449

<212> PRT

<213> Virus influenza

<400> 12

Met Asn Pro Asn Gln Lys Ile Ile Thr Ile Gly Ser Ile Cys Met Val  
 1 5 10 15  
 Thr Gly Ile Val Ser Leu Met Leu Gln Ile Gly Asn Met Ile Ser Ile  
 20 25 30  
 Trp Val Ser His Ser Ile His Thr Gly Asn Gln His Gln Ser Glu Pro  
 35 40 45  
 Ile Ser Asn Thr Asn Phe Leu Thr Glu Lys Ala Val Ala Ser Val Lys  
 50 55 60  
 Leu Ala Gly Asn Ser Ser Leu Cys Pro Ile Asn Gly Trp Ala Val Tyr  
 65 70 75 80  
 Ser Lys Asp Asn Ser Ile Arg Ile Gly Ser Lys Gly Asp Val Phe Val  
 85 90 95  
 Ile Arg Glu Pro Phe Ile Ser Cys Ser His Leu Glu Cys Arg Thr Phe  
 100 105 110  
 Phe Leu Thr Gln Gly Ala Leu Leu Asn Asp Lys His Ser Asn Gly Thr  
 115 120 125  
 Val Lys Asp Arg Ser Pro His Arg Thr Leu Met Ser Cys Pro Val Gly  
 130 135 140  
 Glu Ala Pro Ser Pro Tyr Asn Ser Arg Phe Glu Ser Val Ala Trp Ser  
 145 150 155 160  
 Ala Ser Ala Cys His Asp Gly Thr Ser Trp Leu Thr Ile Gly Ile Ser  
 165 170 175  
 Gly Pro Asp Asn Gly Ala Val Ala Val Leu Lys Tyr Asn Gly Ile Ile  
 180 185 190  
 Thr Asp Thr Ile Lys Ser Trp Arg Asn Asn Ile Leu Arg Thr Gln Glu  
 195 200 205  
 Ser Glu Cys Ala Cys Val Asn Gly Ser Cys Phe Thr Val Met Thr Asp  
 210 215 220  
 Gly Pro Ser Asn Gly Gln Ala Ser His Lys Ile Phe Lys Met Glu Lys  
 225 230 235 240



Gly Lys Val Val Lys Ser Val Glu Leu Asp Ala Pro Asn Tyr His Tyr  
245 250 255

Glu Glu Cys Ser Cys Tyr Pro Asn Ala Gly Glu Ile Thr Cys Val Cys  
260 265 270

Arg Asp Asn Trp His Gly Ser Asn Arg Pro Trp Val Ser Phe Asn Gln  
275 280 285

Asn Leu Glu Tyr Gln Ile Gly Tyr Ile Cys Ser Gly Val Phe Gly Asp  
290 295 300

```

Asn Pro Arg Pro Asn Asp Gly Thr Gly Ser Cys Gly Pro Val Ser Ser
305                      310                      315                      320

```

Asn Gly Ala Tyr Gly Val Lys Gly Phe Ser Phe Lys Tyr Gly Asn Gly  
325 330 335

Val Trp Ile Gly Arg Thr Lys Ser Thr Asn Ser Arg Ser Gly Phe Glu  
340 345 350

Met Ile Trp Asp Pro Asn Gly Trp Thr Glu Thr Asp Ser Ser Phe Ser  
355 360 365

Val Lys Gln Asp Ile Val Ala Ile Thr Asp Trp Ser Gly Tyr Ser Gly  
370 375 380

Ser Phe Val Gln His Pro Glu Leu Thr Gly Leu Asp Cys Ile Arg Pro  
385 390 395 400

Cys Phe Trp Val Glu Leu Ile Arg Gly Arg Pro Lys Glu Ser Thr Ile  
405 410 415

Trp Thr Ser Gly Ser Ser Ile Ser Phe Cys Gly Val Asn Ser Asp Thr  
420 425 430

Val Gly Trp Ser Trp Pro Asp Gly Ala Glu Leu Pro Phe Thr Ile Asp  
435 440 445

**Lys**

<210> 13

<211> 564

<212> PRT

5 <213> Virus influenza

<400> 13

Met Glu Lys Ile Val Leu Leu Phe Ala Ile Val Ser Leu Val Lys Ser

|                                                                 |     |     |     |
|-----------------------------------------------------------------|-----|-----|-----|
| 1                                                               | 5   | 10  | 15  |
| Asp Gln Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Glu Gln Val | 20  | 25  | 30  |
| Asp Thr Ile Met Glu Lys Asn Val Thr Val Thr His Ala Gln Asp Ile | 35  | 40  | 45  |
| Leu Glu Lys Thr His Asn Gly Lys Leu Cys Asp Leu Asp Gly Val Lys | 50  | 55  | 60  |
| Pro Leu Ile Leu Arg Asp Cys Ser Val Ala Gly Trp Leu Leu Gly Asn | 65  | 70  | 75  |
| Pro Met Cys Asp Glu Phe Ile Asn Val Pro Glu Trp Ser Tyr Ile Val | 85  | 90  | 95  |
| Glu Lys Ala Asn Pro Ala Asn Asp Leu Cys Tyr Pro Gly Asp Phe Asn | 100 | 105 | 110 |
| Asp Tyr Glu Glu Leu Lys His Leu Leu Ser Arg Ile Asn His Phe Glu | 115 | 120 | 125 |
| Lys Ile Gln Ile Ile Pro Lys Asn Ser Trp Ser Ser His Glu Ala Ser | 130 | 135 | 140 |
| Leu Gly Val Ser Ser Ala Cys Pro Tyr Gln Gly Lys Ser Ser Phe Phe | 145 | 150 | 155 |
| Arg Asn Val Val Trp Leu Ile Lys Lys Asn Asn Ala Tyr Pro Thr Ile | 165 | 170 | 175 |
| Lys Arg Ser Tyr Asn Asn Thr Asn Gln Glu Asp Leu Leu Val Leu Trp | 180 | 185 | 190 |
| Gly Ile His His Pro Asn Asp Ala Ala Glu Gln Thr Arg Leu Tyr Gln | 195 | 200 | 205 |
| Asn Pro Thr Thr Tyr Ile Ser Val Gly Thr Ser Thr Leu Asn Gln Arg | 210 | 215 | 220 |
| Leu Val Pro Lys Ile Ala Thr Arg Ser Lys Val Asn Gly Gln Asn Gly | 225 | 230 | 235 |
| Arg Met Glu Phe Phe Trp Thr Ile Leu Lys Pro Asn Asp Ala Ile Asn | 245 | 250 | 255 |
| Phe Glu Ser Asn Gly Asn Phe Ile Ala Pro Glu Tyr Ala Tyr Lys Ile | 260 | 265 | 270 |

Val Lys Lys Gly Asp Ser Ala Ile Met Lys Ser Glu Leu Glu Tyr Gly  
 275 280 285  
 Asn Cys Asn Thr Lys Cys Gln Thr Pro Met Gly Ala Ile Asn Ser Ser  
 290 295 300  
 Met Pro Phe His Asn Ile His Pro Leu Thr Ile Gly Glu Cys Pro Lys  
 305 310 315 320  
 Tyr Val Lys Ser Asn Arg Leu Val Leu Ala Thr Gly Leu Arg Asn Ser  
 325 330 335  
 Pro Gln Arg Glu Thr Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile  
 340 345 350  
 Glu Gly Gly Trp Gln Gly Met Val Asp Gly Trp Tyr Gly Tyr His His  
 355 360 365  
 Ser Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Lys Glu Ser Thr Gln  
 370 375 380  
 Lys Ala Ile Asp Gly Val Thr Asn Lys Val Asn Ser Ile Ile Asp Lys  
 385 390 395 400  
 Met Asn Thr Gln Phe Glu Ala Val Gly Arg Glu Phe Asn Asn Leu Glu  
 405 410 415  
 Arg Arg Ile Glu Asn Leu Asn Lys Lys Met Glu Asp Gly Phe Leu Asp  
 420 425 430  
 Val Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Met Glu Asn Glu Arg  
 435 440 445  
 Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu Tyr Asp Lys Val  
 450 455 460  
 Arg Leu Gln Leu Arg Asp Asn Ala Lys Glu Leu Gly Asn Gly Cys Phe  
 465 470 475 480  
 Glu Phe Tyr His Lys Cys Asp Asn Glu Cys Met Glu Ser Val Arg Asn  
 485 490 495  
 Gly Thr Tyr Asp Tyr Pro Gln Tyr Ser Glu Glu Ala Arg Leu Lys Arg  
 500 505 510  
 Glu Glu Ile Ser Gly Val Lys Leu Glu Ser Ile Gly Thr Tyr Gln Ile  
 515 520 525  
 Leu Ser Ile Tyr Ser Thr Val Ala Ser Ser Leu Ala Leu Ala Ile Met  
 530 535 540  
 Val Ala Gly Leu Ser Leu Trp Met Cys Ser Asn Gly Ser Leu Gln Cys  
 545 550 555 560

Arg Ile Cys Ile

<210> 14

<211> 469

5 <212> PRT

&lt;213&gt; Virus influenza

&lt;400&gt; 14

```

Met Asn Pro Asn Gln Lys Ile Thr Thr Ile Gly Ser Ile Cys Met Val
1          5          10          15

Ile Gly Ile Val Ser Leu Met Leu Gln Ile Gly Asn Ile Ile Ser Ile
          20          25          30

Trp Val Ser His Ser Ile Gln Thr Gly Asn Gln His Gln Ala Glu Pro
          35          40          45

Cys Asn Gln Ser Ile Ile Thr Tyr Glu Asn Asn Thr Trp Val Asn Gln
50          55          60

Thr Tyr Val Asn Ile Ser Asn Thr Asn Phe Leu Thr Glu Lys Ala Val
65          70          75          80

Ala Ser Val Thr Leu Ala Gly Asn Ser Ser Leu Cys Pro Ile Ser Gly
          85          90          95

Trp Ala Val Tyr Ser Lys Asp Asn Gly Ile Arg Ile Gly Ser Lys Gly
          100          105          110

Asp Val Phe Val Ile Arg Glu Pro Phe Ile Ser Cys Ser His Leu Glu
          115          120          125

Cys Arg Thr Phe Phe Leu Thr Gln Gly Ala Leu Leu Asn Asp Lys His
          130          135          140

Ser Asn Gly Thr Val Lys Asp Arg Ser Pro His Arg Thr Leu Met Ser
145          150          155          160

Cys Pro Val Gly Glu Ala Pro Ser Pro Tyr Asn Ser Arg Phe Glu Ser
          165          170          175

Val Ala Trp Ser Ala Ser Ala Cys His Asp Gly Thr Ser Trp Leu Thr

```

| 180                                                                                | 185 | 190 |
|------------------------------------------------------------------------------------|-----|-----|
| Ile Gly Ile Ser Gly Pro Asp Asn Gly Ala Val Ala Val Leu Lys Tyr<br>195 200 205     |     |     |
| Asn Gly Ile Ile Thr Asp Thr Ile Lys Ser Trp Arg Asn Asn Ile Met<br>210 215 220     |     |     |
| Arg Thr Gln Glu Ser Glu Cys Ala Cys Val Asn Gly Ser Cys Phe Thr<br>225 230 235 240 |     |     |
| Val Met Thr Asp Gly Pro Ser Asn Gly Gln Ala Ser Tyr Lys Ile Phe<br>245 250 255     |     |     |
| Arg Ile Glu Lys Gly Lys Val Val Lys Ser Ala Glu Leu Asn Ala Pro<br>260 265 270     |     |     |
| Asn Tyr His Tyr Glu Glu Cys Ser Cys Tyr Pro Asp Ala Gly Glu Ile<br>275 280 285     |     |     |
| Thr Cys Val Cys Arg Asp Asn Trp His Gly Ser Asn Arg Pro Trp Val<br>290 295 300     |     |     |
| Ser Phe Asn Gln Asn Leu Glu Tyr Arg Ile Gly Tyr Ile Cys Ser Gly<br>305 310 315 320 |     |     |
| Val Phe Gly Asp Asn Pro Arg Pro Asn Asp Gly Thr Gly Ser Cys Gly<br>325 330 335     |     |     |
| Pro Val Ser Pro Lys Gly Ala Tyr Gly Ile Lys Gly Phe Ser Phe Lys<br>340 345 350     |     |     |
| Tyr Gly Asn Gly Val Trp Ile Gly Arg Thr Lys Ser Thr Asn Ser Arg<br>355 360 365     |     |     |
| Ser Gly Phe Glu Met Ile Trp Asp Pro Asn Gly Trp Thr Gly Thr Asp<br>370 375 380     |     |     |
| Ser Asn Phe Ser Val Lys Gln Asp Ile Val Ala Ile Thr Asp Trp Ser<br>385 390 395 400 |     |     |
| Gly Tyr Ser Gly Ser Phe Val Gln His Pro Glu Leu Thr Gly Leu Asp<br>405 410 415     |     |     |
| Cys Ile Arg Pro Cys Phe Trp Val Glu Leu Ile Arg Gly Arg Pro Lys<br>420 425 430     |     |     |
| Glu Ser Thr Ile Trp Thr Ser Gly Ser Ser Ile Ser Phe Cys Gly Val<br>435 440 445     |     |     |
| Asn Ser Asp Thr Val Gly Trp Ser Trp Pro Asp Gly Ala Glu Leu Pro<br>450 455 460     |     |     |

Phe Thr Ile Asp Lys  
465

<210> 15

<211> 564

<212> PRT

<213> Virus influenza

<400> 15

Met Glu Lys Ile Val Leu Leu Leu Ala Thr Val Ser Leu Val Lys Ser  
 1 5 10 15  
 Asp Gln Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Glu Gln Val  
 20 25 30  
 Asp Thr Ile Met Glu Lys Asn Val Thr Val Thr His Ala Gln Asp Ile  
 35 40 45  
 Leu Glu Arg Thr His Asn Gly Lys Leu Cys Asp Leu Asn Gly Val Lys  
 50 55 60  
 Pro Leu Ile Leu Arg Asp Cys Ser Val Ala Gly Trp Leu Leu Gly Asn  
 65 70 75 80  
 Pro Met Cys Asp Glu Phe Ile Asn Val Pro Glu Trp Ser Tyr Ile Val  
 85 90 95  
 Glu Lys Ala Ser Pro Ala Asn Asp Leu Cys Tyr Pro Gly Asn Phe Asn  
 100 105 110  
 Asp Tyr Glu Glu Leu Lys His Leu Leu Ser Arg Ile Asn His Phe Glu  
 115 120 125  
 Lys Ile Gln Ile Ile Pro Lys Ser Ser Trp Ser Asn His Asp Ala Ser  
 130 135 140  
 Ser Gly Val Ser Ser Ala Cys Pro Tyr Leu Gly Arg Ser Ser Phe Phe  
 145 150 155 160  
 Arg Asn Val Val Trp Leu Ile Lys Lys Asn Ser Ser Tyr Pro Thr Ile  
 165 170 175  
 Lys Arg Ser Tyr Asn Asn Thr Asn Gln Glu Asp Leu Leu Val Leu Trp  
 180 185 190

Gly Ile His His Pro Asn Asp Ala Ala Glu Gln Thr Arg Leu Tyr Gln  
 195 200 205

Asn Pro Thr Thr Tyr Ile Ser Val Gly Thr Ser Thr Leu Asn Gln Arg  
 210 215 220

Leu Val Pro Glu Ile Ala Thr Arg Pro Lys Val Asn Gly Gln Ser Gly  
 225 230 235 240

Arg Met Glu Phe Phe Trp Thr Ile Leu Lys Pro Asn Asp Ala Ile Asn  
 245 250 255

Phe Glu Ser Asn Gly Asn Phe Ile Ala Pro Glu Tyr Ala Tyr Lys Ile  
 260 265 270

Val Lys Lys Gly Asp Ser Thr Ile Met Lys Ser Glu Leu Glu Tyr Gly  
 275 280 285

Asn Cys Asn Thr Lys Cys Gln Thr Pro Met Gly Ala Ile Asn Ser Ser  
 290 295 300

Met Pro Phe His Asn Ile His Pro Leu Thr Ile Gly Glu Cys Pro Lys  
 305 310 315 320

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

Pro Gln Arg Glu Thr Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile  
 340 345 350

Glu Gly Gly Trp Gln Gly Met Val Asp Gly Trp Tyr Gly Tyr His His  
 355 360 365

Ser Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Gln Glu Ser Thr Gln  
 370 375 380

Lys Ala Ile Asp Gly Val Thr Asn Lys Val Asn Ser Ile Ile Asn Lys  
 385 390 395 400

Met Asn Thr Gln Phe Glu Ala Val Gly Arg Glu Phe Asn Asn Leu Glu  
 405 410 415

Arg Arg Ile Glu Asn Leu Asn Lys Lys Met Glu Asp Gly Phe Leu Asp  
 420 425 430

Val Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Met Glu Asn Glu Arg  
 435 440 445

Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu Tyr Asp Lys Val

450

455

460

Arg Leu Gln Leu Arg Asp Asn Ala Lys Glu Leu Gly Asn Gly Cys Phe  
465 470 475 480

Glu Phe Tyr His Lys Cys Asp Asn Glu Cys Met Glu Ser Val Lys Asn  
485 490 495

Gly Thr Tyr Asp Tyr Pro Gln Tyr Ser Glu Glu Ala Arg Leu Asn Arg  
500 505 510

Glu Glu Ile Ser Gly Val Lys Leu Glu Ser Met Gly Thr Tyr Gln Ile  
515 520 525

Leu Ser Ile Tyr Ser Thr Val Ala Ser Ser Leu Ala Leu Ala Ile Met  
530 535 540

Val Ala Gly Leu Ser Leu Trp Met Cys Ser Asn Gly Ser Leu Gln Cys  
545 550 555 560

Arg Ile Cys Ile

<210> 16

<211> 450

<212> PRT

5 <213> Virus influenza

<400> 16

Met Asn Pro Asn Gln Lys Ile Ile Thr Ile Gly Ser Ile Cys Met Val  
1 5 10 15

Val Gly Ile Ile Ser Leu Met Leu Gln Ile Gly Asn Thr Ile Ser Val  
20 25 30

Trp Val Ser His Ile Ile Lys Thr Trp His Pro Asn Gln Pro Glu Pro  
35 40 45

Cys Asn Gln Ser Ile Asn Phe Tyr Thr Glu Gln Ala Ala Ala Ser Val  
50 55 60

Thr Leu Ala Gly Asn Ser Ser Leu Cys Pro Ile Ser Gly Trp Ala Ile  
65 70 75 80

Tyr Ser Lys Asp Asn Ser Ile Arg Ile Gly Ser Lys Gly Asp Val Phe  
85 90 95

Val Ile Arg Glu Pro Phe Ile Ser Cys Ser His Leu Glu Cys Arg Thr  
100 105 110



Phe Phe Leu Thr Gln Gly Ala Leu Leu Asn Asp Lys His Ser Asn Gly  
 115 120 125

Thr Val Lys Asp Arg Ser Pro Tyr Arg Thr Leu Met Ser Cys Pro Val  
 130 135 140

Gly Glu Ala Pro Ser Pro Tyr Asn Ser Arg Phe Glu Ser Val Ala Trp  
 145 150 155 160

Ser Ala Ser Ala Cys His Asp Gly Ile Ser Trp Leu Thr Ile Gly Ile  
 165 170 175

Ser Gly Pro Asp Asn Gly Ala Val Ala Val Leu Lys Tyr Asn Gly Ile  
 180 185 190

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

Glu Ser Glu Cys Ala Cys Val Asn Gly Ser Cys Phe Thr Val Met Thr  
 210 215 220

Asp Gly Pro Ser Asn Glu Gln Ala Ser Tyr Lys Ile Phe Lys Ile Glu  
 225 230 235 240

Lys Gly Arg Val Val Lys Ser Val Glu Leu Asn Ala Pro Asn Tyr His  
 245 250 255

Tyr Glu Glu Cys Ser Cys Tyr Pro Asp Ala Gly Glu Ile Thr Cys Val  
 260 265 270

Cys Arg Asp Asn Trp His Gly Ser Asn Arg Pro Trp Val Ser Phe Asn  
 275 280 285

Gln Asn Leu Glu Tyr Gln Ile Gly Tyr Ile Cys Ser Gly Val Phe Gly  
 290 295 300

Asp Ser Pro Arg Pro Asn Asp Gly Thr Gly Ser Cys Gly Pro Val Ser  
 305 310 315 320

Leu Asn Gly Ala Tyr Gly Val Lys Gly Phe Ser Phe Lys Tyr Gly Asn  
 325 330 335

Gly Val Trp Ile Gly Arg Thr Lys Ser Thr Ser Ser Arg Ser Gly Phe  
 340 345 350

Glu Met Ile Trp Asp Pro Asn Gly Trp Thr Glu Thr Asp Ser Ser Phe  
 355 360 365

Ser Leu Lys Gln Asp Ile Ile Ala Ile Thr Asp Trp Ser Gly Tyr Ser  
370 375 380

Gly Ser Phe Ile Gln His Pro Glu Leu Thr Gly Leu Asn Cys Met Arg  
385 390 395 400

Pro Cys Phe Trp Val Glu Leu Ile Arg Gly Arg Pro Lys Glu Lys Thr  
405 410 415

Ile Trp Thr Ser Gly Ser Ser Ile Ser Phe Cys Gly Val Asn Ser Asp  
420 425 430

Thr Val Gly Trp Ser Trp Pro Asp Gly Ala Glu Leu Pro Tyr Thr Ile  
435 440 445

Asp Lys  
450

<210> 17

<211> 564

<212> PRT

5 <213> Virus influenza

<400> 17

Met Glu Lys Ile Val Leu Leu Leu Ala Thr Val Ser Leu Val Lys Ser  
1 5 10 15

Asp Gln Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Glu Gln Val  
20 25 30

Asp Thr Ile Met Glu Lys Asn Val Thr Val Thr His Ala Gln Asp Ile  
35 40 45

Leu Glu Arg Thr His Asn Gly Lys Leu Cys Asp Leu Asn Gly Val Lys  
50 55 60

Pro Leu Ile Leu Arg Asp Cys Ser Val Ala Gly Trp Leu Leu Gly Asn  
65 70 75 80

Pro Met Cys Asp Glu Phe Ile Asn Val Pro Glu Trp Ser Tyr Ile Val  
85 90 95

Glu Lys Ala Ser Pro Ala Asn Asp Leu Cys Tyr Pro Gly Asn Phe Asn  
100 105 110

Asp Tyr Glu Glu Leu Lys His Leu Leu Ser Arg Ile Asn His Phe Glu  
115 120 125

Lys Ile Gln Ile Ile Pro Lys Ser Ser Trp Ser Asn His Asp Ala Ser  
130 135 140

Ser Gly Val Ser Ser Ala Cys Pro Tyr Leu Gly Arg Ser Ser Phe Phe  
 145 150 155 160

Arg Asn Val Val Trp Leu Ile Lys Lys Asn Ser Ser Tyr Pro Thr Ile  
 165 170 175

Lys Arg Ser Tyr Asn Asn Thr Asn Gln Glu Asp Leu Leu Val Leu Trp  
 180 185 190

Gly Ile His His Pro Asn Asp Ala Ala Glu Gln Thr Arg Leu Tyr Gln  
 195 200 205

Asn Pro Thr Thr Tyr Ile Ser Val Gly Thr Ser Thr Leu Asn Gln Arg  
 210 215 220

Leu Val Ser Glu Ile Ala Thr Arg Pro Lys Val Asn Gly Gln Ser Gly  
 225 230 235 240

Arg Met Glu Phe Phe Trp Thr Ile Leu Lys Pro Asn Asp Ala Ile Asn  
 245 250 255

Phe Glu Ser Asn Gly Asn Phe Ile Ala Pro Glu Tyr Ala Tyr Lys Ile  
 260 265 270

Val Lys Lys Gly Asp Ser Thr Ile Met Lys Ser Glu Leu Glu Tyr Gly  
 275 280 285

Asn Cys Asn Thr Lys Cys Gln Thr Pro Met Gly Ala Ile Asn Ser Ser  
 290 295 300

Met Pro Phe His Asn Ile His Pro Leu Thr Ile Gly Glu Cys Pro Lys  
 305 310 315 320

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

Pro Gln Arg Glu Thr Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile  
 340 345 350

Glu Gly Gly Trp Gln Gly Met Val Asp Gly Trp Tyr Gly Tyr His His  
 355 360 365

Ser Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Gln Glu Ser Thr Gln  
 370 375 380

Lys Ala Ile Asp Gly Val Thr Asn Lys Val Asn Ser Ile Ile Asn Lys  
 385 390 395 400

Met Asn Thr Gln Phe Glu Ala Val Gly Arg Glu Phe Asn Asn Leu Glu  
 405 410 415

Arg Arg Ile Glu Asn Leu Asn Lys Lys Met Glu Asp Gly Phe Leu Asp  
 420 425 430

Val Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Met Glu Asn Glu Arg  
 435 440 445

Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu Tyr Asp Lys Val  
 450 455 460

Arg Leu Gln Leu Arg Asp Asn Ala Lys Glu Leu Gly Asn Gly Cys Phe  
 465 470 475 480

Glu Phe Tyr His Lys Cys Asp Asn Glu Cys Met Glu Ser Val Lys Asn  
 485 490 495

Gly Thr Tyr Asp Tyr Pro Gln Tyr Ser Glu Glu Ala Arg Leu Asn Arg  
 500 505 510

Glu Glu Ile Ser Gly Val Lys Leu Glu Ser Met Gly Thr Tyr Gln Ile  
 515 520 525

Leu Ser Ile Tyr Ser Thr Val Ala Ser Ser Leu Ala Leu Ala Ile Met  
 530 535 540

Val Ala Gly Leu Ser Leu Trp Met Cys Ser Asn Gly Ser Leu Gln Cys  
 545 550 555 560

Arg Ile Cys Ile

<210> 18

<211> 450

<212> PRT

5 <213> Virus influenza

<400> 18

Met Asn Pro Asn Gln Lys Ile Ile Thr Ile Gly Ser Ile Cys Met Val  
 1 5 10 15

Val Gly Ile Ile Ser Leu Met Leu Gln Ile Gly Asn Thr Ile Ser Val  
 20 25 30

Trp Val Ser His Ile Ile Lys Thr Trp His Pro Asn Gln Pro Glu Pro  
 35 40 45

Cys Asn Gln Ser Ile Asn Phe Tyr Thr Glu Gln Ala Ala Ala Ser Val

| 50                                                                                 | 55 | 60 |
|------------------------------------------------------------------------------------|----|----|
| Thr Leu Ala Gly Asn Ser Ser Leu Cys Pro Ile Ser Gly Trp Ala Ile<br>65 70 75 80     |    |    |
| Tyr Ser Lys Asp Asn Ser Ile Arg Ile Gly Ser Lys Gly Asp Val Phe<br>85 90 95        |    |    |
| Val Ile Arg Glu Pro Phe Ile Ser Cys Ser His Leu Glu Cys Arg Thr<br>100 105 110     |    |    |
| Phe Phe Leu Thr Gln Gly Ala Leu Leu Asn Asp Lys His Ser Asn Gly<br>115 120 125     |    |    |
| Thr Val Lys Asp Arg Ser Pro Tyr Arg Thr Leu Met Ser Cys Pro Val<br>130 135 140     |    |    |
| Gly Glu Ala Pro Ser Pro Tyr Asn Ser Arg Phe Glu Ser Val Ala Trp<br>145 150 155 160 |    |    |
| Ser Ala Ser Ala Cys His Asp Gly Ile Ser Trp Leu Thr Ile Gly Ile<br>165 170 175     |    |    |
| Ser Gly Pro Asp Asn Gly Ala Val Ala Val Leu Lys Tyr Asn Gly Ile<br>180 185 190     |    |    |
| Ile Thr Asp Thr Ile Lys Ser Trp Arg Asn Asn Thr Leu Arg Thr Gln<br>195 200 205     |    |    |
| Glu Ser Glu Cys Ala Cys Val Asn Gly Ser Cys Phe Thr Val Met Thr<br>210 215 220     |    |    |
| Asp Gly Pro Ser Asn Glu Gln Ala Ser Tyr Lys Ile Phe Lys Ile Glu<br>225 230 235 240 |    |    |
| Lys Gly Arg Val Val Lys Ser Val Glu Leu Asn Ala Pro Asn Tyr His<br>245 250 255     |    |    |
| Tyr Glu Glu Cys Ser Cys Tyr Pro Asp Ala Gly Glu Ile Thr Cys Val<br>260 265 270     |    |    |
| Cys Arg Asp Asn Trp His Gly Ser Asn Arg Pro Trp Val Ser Phe Asn<br>275 280 285     |    |    |
| Gln Asn Leu Glu Tyr Gln Ile Gly Tyr Ile Cys Ser Gly Val Phe Gly<br>290 295 300     |    |    |
| Asp Ser Pro Arg Pro Asn Asp Gly Thr Gly Ser Cys Gly Pro Val Ser<br>305 310 315 320 |    |    |

Leu Asn Gly Ala Tyr Gly Val Lys Gly Phe Ser Phe Lys Tyr Gly Asn  
                   325                  330                  335

Gly Val Trp Ile Gly Arg Thr Lys Ser Thr Ser Ser Arg Ser Gly Phe  
                   340                  345                  350

Glu Met Ile Trp Asp Pro Asn Gly Trp Thr Glu Thr Asp Ser Ser Phe  
                   355                  360                  365

Ser Leu Lys Gln Asp Ile Ile Ala Ile Thr Asp Trp Ser Gly Tyr Ser  
           370                  375                  380

Gly Ser Phe Ile Gln His Pro Glu Leu Thr Gly Leu Asn Cys Met Arg  
   385                  390                  395                  400

Pro Cys Phe Trp Val Glu Leu Ile Arg Gly Arg Pro Lys Glu Lys Thr  
                   405                  410                  415

Ile Trp Thr Ser Gly Ser Ser Ile Ser Phe Cys Gly Val Asn Ser Asp  
                   420                  425                  430

Thr Val Gly Trp Ser Trp Pro Asp Gly Ala Glu Leu Pro Tyr Thr Ile  
           435                  440                  445

Asp Lys  
   450

<210> 19

<211> 558

<212> PRT

5 <213> Virus influenza

<400> 19

Met Glu Ala Ile Pro Leu Ile Thr Ile Leu Leu Val Val Thr Ala Ser  
   1                  5                  10                  15

Asn Ala Asp Lys Ile Cys Ile Gly Tyr Gln Ser Thr Asn Ser Thr Glu  
           20                  25                  30

Thr Val Asp Thr Leu Thr Glu Asn Asn Val Pro Val Thr His Ala Lys  
           35                  40                  45

Glu Leu Leu His Thr Glu His Asn Gly Met Leu Cys Ala Thr Asn Leu  
   50                  55                  60

Gly Arg Pro Leu Ile Leu Asp Thr Cys Thr Ile Glu Gly Leu Ile Tyr  
   65                  70                  75                  80

Gly Asn Pro Ser Cys Asp Leu Leu Leu Gly Gly Arg Glu Trp Ser Tyr  
 85 90 95  
 Ile Val Glu Arg Pro Ser Ala Val Asn Gly Met Cys Tyr Pro Gly Asn  
 100 105 110  
 Val Glu Asn Leu Glu Glu Leu Arg Ser Phe Phe Ser Ser Ala Ser Ser  
 115 120 125  
 Tyr Gln Arg Ile Gln Ile Phe Pro Asp Thr Ile Trp Asn Val Ser Tyr  
 130 135 140  
 Ser Gly Thr Ser Lys Ala Cys Ser Asp Ser Phe Tyr Arg Ser Met Arg  
 145 150 155 160  
 Trp Leu Thr Gln Lys Asn Asn Ala Tyr Pro Ile Gln Asp Ala Gln Tyr  
 165 170 175  
 Thr Asn Asn Arg Gly Lys Ser Ile Leu Phe Met Trp Gly Ile Asn His  
 180 185 190  
 Pro Pro Thr Asp Thr Ala Gln Thr Asn Leu Tyr Thr Arg Thr Asp Thr  
 195 200 205  
 Thr Thr Ser Val Ala Thr Glu Asp Ile Asn Arg Thr Phe Lys Pro Val  
 210 215 220  
 Ile Gly Pro Arg Pro Leu Val Asn Gly Leu Gln Gly Arg Ile Asp Tyr  
 225 230 235 240  
 Tyr Trp Ser Val Leu Lys Pro Gly Gln Thr Leu Arg Val Arg Ser Asn  
 245 250 255  
 Gly Asn Leu Ile Ala Pro Trp Tyr Gly His Ile Leu Ser Gly Glu Ser  
 260 265 270  
 His Gly Arg Ile Leu Lys Thr Asp Leu Asn Ser Gly Ser Cys Val Val  
 275 280 285  
 Gln Cys Gln Thr Glu Arg Gly Gly Leu Asn Thr Thr Leu Pro Phe His  
 290 295 300  
 Asn Val Ser Lys Tyr Ala Phe Gly Asn Cys Pro Lys Tyr Val Gly Val  
 305 310 315 320  
 Lys Ser Leu Lys Leu Ala Val Gly Leu Arg Asn Val Pro Ala Arg Ser  
 325 330 335  
 Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly Gly Trp

340                      345                      350  
 Ser Gly Leu Val Ala Gly Trp Tyr Gly Phe Gln His Ser Asn Asp Gln  
           355                      360                      365  
 Gly Val Gly Ile Ala Ala Asp Arg Asp Ser Thr Gln Arg Ala Ile Asp  
           370                      375                      380  
 Lys Ile Thr Ser Lys Val Asn Asn Ile Val Asp Lys Met Asn Lys Gln  
 385                      390                      395                      400  
 Tyr Glu Ile Ile Asp His Glu Phe Ser Glu Val Glu Asn Arg Leu Asn  
                           405                      410                      415  
 Met Ile Asn Asn Lys Ile Asp Asp Gln Ile Gln Asp Ile Trp Ala Tyr  
                           420                      425                      430  
 Asn Ala Glu Leu Leu Val Leu Leu Glu Asn Gln Lys Thr Leu Asp Glu  
                           435                      440                      445  
 His Asp Ala Asn Val Asn Asn Leu Tyr Asn Lys Val Lys Arg Ala Leu  
                           450                      455                      460  
 Gly Ser Asn Ala Met Glu Asp Gly Lys Gly Cys Phe Glu Leu Tyr His  
 465                      470                      475                      480  
 Lys Cys Asp Asp Gln Cys Met Glu Thr Ile Arg Asn Gly Thr Tyr Asn  
                           485                      490                      495  
 Arg Arg Lys Tyr Lys Glu Glu Ser Arg Leu Glu Arg Gln Lys Ile Glu  
                           500                      505                      510  
 Gly Val Lys Leu Glu Ser Glu Gly Thr Tyr Lys Ile Leu Thr Ile Tyr  
                           515                      520                      525  
 Ser Thr Val Ala Ser Ser Leu Val Ile Ala Met Gly Phe Ala Ala Phe  
                           530                      535                      540  
 Leu Phe Trp Ala Met Ser Asn Gly Ser Cys Arg Cys Asn Ile  
 545                      550                      555  
 <210> 20  
 <211> 469  
 <212> PRT  
 5 <213> Virus influenza  
 <400> 20  
 Met Asn Pro Asn Gln Lys Ile Ile Ala Ile Gly Ser Val Ser Leu Thr  
 1                      5                      10                      15



Ile Ala Thr Ile Cys Leu Leu Met Gln Ile Ala Ile Leu Ala Thr Thr  
 20 25 30  
 Met Thr Leu His Phe Lys Gln Asn Glu Cys Ile Asn Ser Ser Asn Asn  
 35 40 45  
 Gln Val Val Pro Cys Glu Pro Ile Ile Ile Glu Arg Asn Ile Thr Glu  
 50 55 60  
 Ile Val His Leu Asn Ser Thr Thr Leu Glu Lys Glu Ile Cys Pro Lys  
 65 70 75 80  
 Val Ala Asp Tyr Arg Asn Trp Ser Lys Pro Gln Cys Gln Ile Thr Gly  
 85 90 95  
 Phe Ala Pro Phe Ser Lys Asp Asn Ser Ile Arg Leu Ser Ala Gly Gly  
 100 105 110  
 Asp Ile Trp Val Thr Arg Glu Pro Tyr Val Ser Cys Gly Leu Gly Lys  
 115 120 125  
 Cys Tyr Gln Phe Ala Leu Gly Gln Gly Thr Thr Leu Glu Asn Lys His  
 130 135 140  
 Ser Asn Gly Thr Ala His Asp Arg Thr Pro His Arg Thr Leu Leu Met  
 145 150 155 160  
 Asn Glu Leu Gly Val Pro Phe His Leu Ala Thr Lys Gln Val Cys Ile  
 165 170 175  
 Ala Trp Ser Ser Ser Ser Cys His Asp Gly Lys Ala Trp Leu His Val  
 180 185 190  
 Cys Val Thr Gly Asp Asp Arg Asn Ala Thr Ala Ser Ile Ile Tyr Asp  
 195 200 205  
 Gly Ile Leu Val Asp Ser Ile Gly Ser Trp Ser Lys Asn Ile Leu Arg  
 210 215 220  
 Thr Gln Glu Ser Glu Cys Val Cys Ile Asn Gly Thr Cys Ala Val Val  
 225 230 235 240  
 Met Thr Asp Gly Ser Ala Ser Gly Arg Ala Asp Thr Arg Ile Leu Phe  
 245 250 255  
 Ile Arg Glu Gly Lys Ile Ala His Ile Ser Pro Leu Ser Gly Ser Ala  
 260 265 270

Gln His Val Glu Glu Cys Ser Cys Tyr Pro Arg Tyr Pro Glu Val Arg  
275 280 285

Cys Val Cys Arg Asp Asn Trp Lys Gly Ser Asn Arg Pro Val Leu Tyr  
290 295 300

Ile Asn Met Ala Asn Tyr Ser Ile Asp Ser Ser Tyr Val Cys Ser Gly  
305 310 315 320

Leu Val Gly Asp Thr Pro Arg Asn Asp Asp Arg Ser Ser Ser Ser Asn  
325 330 335

Cys Arg Asp Pro Asn Asn Glu Arg Gly Ala Pro Gly Val Lys Gly Trp  
340 345 350

Ala Phe Asp Asn Gly Asn Asp Ile Trp Met Gly Arg Thr Ile Lys Lys  
355 360 365

Asp Ser Arg Ser Gly Tyr Glu Thr Phe Arg Val Ile Gly Gly Trp Thr  
370 375 380

Thr Ala Asn Ser Lys Ser Gln Ile Asn Arg Gln Val Ile Val Asp Ser  
385 390 395 400

Asp Asn Ser Ser Gly Tyr Ser Gly Ile Phe Ser Val Glu Gly Lys Ser  
405 410 415

Cys Ile Asn Arg Cys Phe Tyr Val Glu Leu Ile Arg Gly Arg Pro Lys  
420 425 430

Glu Thr Arg Val Trp Trp Thr Ser Asn Ser Ile Ile Val Phe Cys Gly  
435 440 445

Thr Ser Gly Thr Tyr Gly Thr Gly Ser Trp Pro Asp Gly Ala Asn Ile  
450 455 460

Asn Phe Met Pro Ile  
465

La divulgación comprende adicionalmente los puntos siguientes:

- 5 1. Un polipéptido aislado, en el que dicho polipéptido se selecciona del grupo que consiste en:
  - a) un polipéptido codificado por una secuencia de polinucleótidos de las SEC ID N° 1-10;
  - b) un polipéptido de las SEC ID N° 11-20;
  - 10 c) la forma madura del polipéptido de las SEC ID N° 11 -20;
  - d) un polipéptido codificado por una secuencia de polinucleótidos que hibrida en condiciones de rigurosidad elevada con una secuencia de polinucleótidos que codifica (a) (b) o (c); y
  - 15 d) un polipéptido que tiene una identidad de secuencia de al menos un 90% con el polipéptido de (b).
2. Una composición inmunogénica que comprende una cantidad inmunológicamente eficaz de al menos un polipéptido del punto 1.
- 20 3. Un anticuerpo aislado que se une específicamente al polipéptido del punto 1.

4. Un método para estimular el sistema inmunitario de un individuo para que produzca una respuesta inmunitaria protectora contra el virus influenza, comprendiendo el método administrar al individuo una cantidad inmunológicamente eficaz del polipéptido del punto 1 en un vehículo fisiológicamente aceptable.
- 5 5. Un virus influenza recombinante que comprende el polipéptido del punto 1.
6. Una composición inmunogénica que comprende una cantidad inmunológicamente eficaz del virus influenza recombinante del punto 5.
- 10 7. Un método para estimular el sistema inmunitario de un individuo para que produzca una respuesta inmunitaria protectora contra el virus influenza, comprendiendo el método administrar al individuo una cantidad inmunológicamente eficaz del virus influenza recombinante del punto 5 en un vehículo fisiológicamente aceptable.
- 15 8. Un ácido nucleico aislado, en el que dicho ácido nucleico se selecciona del grupo que consiste en:
- a) una secuencia de polinucleótidos de las SEC ID N° 1-10 o una secuencia complementaria de la misma;
- b) una secuencia de polinucleótidos que codifica un polinucleótido de las SEC ID N° 11-20 o una secuencia complementaria de la misma;
- 20 c) una secuencia de polinucleótidos que hibrida en condiciones de rigurosidad elevada con sustancialmente la totalidad de la longitud de la secuencia de polinucleótidos (a); y
- 25 d) una secuencia de polinucleótidos que tiene una identidad de secuencia de al menos un 98% con la secuencia de polinucleótidos de (a).
9. Una composición inmunogénica que comprende al menos uno de los ácidos nucleicos del punto 8.
- 30 10. Una célula que comprende al menos un ácido nucleico del punto 8.
11. Un vector que comprende el ácido nucleico del punto 8.
12. El vector del punto 12, en el que el vector es un plásmido, un cósmico, un fago, un virus o un fragmento de un virus.
- 35 13. El vector del punto 12, en el que el vector es un vector de expresión.
14. Una célula que comprende el vector del punto 13.
- 40 15. Un virus influenza que comprende uno o más ácidos nucleicos del punto 8.
16. El virus del punto 15, en el que el virus es un virus reagrupado.
- 45 17. Un virus influenza con reagrupamiento 6:2, donde dicho virus comprende 6 regiones codificantes de genes de A/Ann Arbor/6/60 y 2 regiones codificantes de genes que codifican un polipéptido seleccionado del grupo que consiste en: los polipéptidos de las SEC ID N° 11-20.
18. Un método para producir un virus influenza recombinante, comprendiendo el método:
- 50 cultivar la célula de punto 14 en un medio de cultivo adecuado en condiciones que permitan la expresión de ácido nucleico; y
- aislar el virus influenza recombinante de una población celular que comprende dicha célula o el medio.
- 55 19. Una composición inmunogénica que comprende una cantidad inmunológicamente eficaz del virus influenza recombinante del punto 17.
20. Un método para estimular el sistema inmunitario de un individuo para que produzca una respuesta inmunitaria protectora contra el virus influenza, comprendiendo el método administrar al individuo una cantidad inmunológicamente eficaz del virus influenza recombinante del punto 17 en un vehículo fisiológicamente eficaz.
- 60 21. Un método para producir un polipéptido aislado o recombinante, comprendiendo el método: cultivar la célula huésped del punto 10 en un medio de cultivo adecuado en condiciones que permitan la expresión de dicho ácido nucleico; y aislar el polipéptido de una o más de las células huésped o el medio.
- 65 22. Un método de tratamiento profiláctico o terapéutico de una infección viral en un sujeto, comprendiendo el

método: administrar al sujeto un virus del punto 17 en una cantidad eficaz para producir una respuesta inmunogénica contra la infección viral.

23. El método del punto 22, en el que el sujeto es un ser humano.

24. La composición inmunogénica del punto 19, en la que la hemaglutinina comprende un sitio de escisión polibásico modificado.

25. Una vacuna con virus influenza vivos atenuados que comprende la composición del punto 19.

26. Una vacuna con virus divididos o muertos que comprende la composición del punto 19.

27. Una vacuna con virus influenza vivos atenuados que comprende la composición del punto 24.

28. Una vacuna con virus divididos o muertos que comprende la composición del punto 24.

29. Un método para producir virus influenza en cultivo celular, comprendiendo el método:

i) introducir en una población de células huésped, población de células huésped que es capaz de soportar replicación del virus influenza, una pluralidad de vectores que comprende ácido nucleico que codifica al menos 6 segmentos internos del genoma de una primera cepa de influenza, en el que la primera cepa de influenza es A/Ann Arbor/6/60; y al menos un segmento genómico que codifica un antígeno de superficie de influenza inmunogénico de una segunda cepa de influenza, en el que dicha segunda cepa es una cepa de influenza pandémica,

ii) cultivar la población de células huésped a una temperatura igual o inferior a 35°C; y

iii) recuperar una pluralidad de virus de influenza.

30. El método del punto 29, en el que la pluralidad de vectores comprende al menos un ácido nucleico aislado, en el que dicho ácido nucleico se selecciona del grupo que consiste en:

a) una secuencia de polinucleótidos de las SEC ID N° 1-10 o una secuencia complementaria de la misma;

b) una secuencia de polinucleótidos que codifica un polipéptido de las SEC ID N° 11-20 o una secuencia de polinucleótidos complementaria de la misma;

c) una secuencia de polinucleótidos que hibrida en condiciones de rigurosidad elevada con sustancialmente la totalidad de la longitud de la secuencia de polinucleótidos (a); y

d) una secuencia de polinucleótidos que tiene una identidad de secuencia de al menos un 98% con la secuencia de polinucleótidos de (a).

31. Una composición inmunogénica que comprende una cantidad inmunológicamente eficaz del virus influenza del punto 29.

32. Una composición inmunogénica que comprende una cantidad inmunológicamente eficaz del virus influenza del punto 30.

33. Un método para estimular el sistema inmunitario de un individuo para que produzca una respuesta inmunitaria protectora contra el virus influenza, comprendiendo el método administrar al individuo una cantidad inmunológicamente eficaz del virus influenza del punto 29 en un vehículo fisiológicamente eficaz.

34. Un método para estimular el sistema inmunitario de un individuo para que produzca una respuesta inmunitaria protectora contra el virus influenza, comprendiendo el método administrar al individuo una cantidad inmunológicamente eficaz del virus influenza del punto 30 en un vehículo fisiológicamente eficaz.

35. Un método para estimular el sistema inmunitario de un individuo para que produzca una respuesta inmunitaria protectora contra el virus influenza, comprendiendo el método administrar al individuo la composición inmunogénica del punto 31.

36. Un método para estimular el sistema inmunitario de un individuo para que produzca una respuesta inmunitaria protectora contra el virus influenza, comprendiendo el método administrar al individuo la composición inmunogénica del punto 32.

37. Una vacuna con virus influenza vivos atenuados que comprende la composición inmunogénica del punto 31.

38. Una vacuna con virus divididos o muertos que comprende la composición inmunogénica del punto 32.

**REIVINDICACIONES**

1. Un virus influenza con reagrupamiento 6:2, en el que dicho virus comprende 6 regiones codificantes de genes que codifican para segmentos de genoma internos de A/Ann Arbor/6/60 y 2 regiones codificantes de genes que codifican un polipéptido de hemaglutinina y un polipéptido de neuraminidasa, en el que el polipéptido de hemaglutinina comprende la secuencia de aminoácidos de SEC ID N° 11 y el polipéptido de neuraminidasa comprende la secuencia de aminoácidos de SEC ID N° 12.
2. Una composición inmunogénica que comprende una cantidad inmunológicamente eficaz del virus influenza reagrupado según la reivindicación 1.
3. Una vacuna contra el virus influenza con virus atenuados vivos que comprende la composición según la reivindicación 2.
4. Una vacuna de virus divididos o virus inactivados que comprende la composición inmunogénica según la reivindicación 2.
5. El virus influenza reagrupado según la reivindicación 1 o la composición inmunogénica según la reivindicación 2, para uso en un método de tratamiento mediante la estimulación del sistema inmunitario de un sujeto para producir una respuesta inmunitaria protectora contra el virus influenza, en el que el virus influenza reagrupado se debe administrar al sujeto en una cantidad inmunológicamente eficaz y en un vehículo fisiológicamente eficaz.
6. El virus influenza reagrupado según la reivindicación 1 o la composición inmunogénica según la reivindicación 2, para uso en un método de tratamiento profiláctico de una infección viral en un sujeto, en el que el virus influenza reagrupado debe administrarse al sujeto en una cantidad eficaz para producir una respuesta inmunogénica contra la infección viral.
7. Un método para producir virus influenza en cultivo celular, comprendiendo el método:
  - i) introducir en una población de células huésped, población de células huésped que puede soportar la replicación de virus influenza, una pluralidad de vectores que comprenden secuencias de ácido nucleico que corresponden a:
    - al menos 6 segmentos internos del genoma de la cepa de influenza A/Ann Arbor/6/60; y
    - al menos 2 regiones codificantes de genes que codifican un polipéptido de hemaglutinina y un polipéptido de neuraminidasa,
 en el que el polipéptido de hemaglutinina comprende la secuencia de aminoácidos de SEC ID N° 11 y el polipéptido de neuraminidasa comprende la secuencia de aminoácidos de SEC ID N° 12;
  - ii) cultivar la población de células huésped en presencia de tripsina a una temperatura inferior o igual a 35°C; y
  - iii) recuperar una pluralidad de virus influenza.

Figura 1

HA de tipo natural de VN/1203/2004

CCT CAA AGA GAG AGA AGA AGA AAG AGA ↓ GGA TTA TTT  
 Pro Gln Arg Glu Arg Arg Arg Lys Lys Arg Gly Leu Phe

HA modificada

CCT CAA AGA GAG ACT CGA ↓ GGA TTA TTT  
 Pro Gln Arg Glu Thr Arg Gly Leu Phe

↓ Sitio de escisión en dominios HA1 y HA2.

Los residuos que se mutagenizaron están subrayados

Figura 2

## Aislamiento de virus de hisopos

| Virus                           | Mortalidad<br>(muertos/total) | Orofaringeos                     |                                                                 | Cloacales                        |                                                                 | Anticuerpos<br>detectados/<br>total |
|---------------------------------|-------------------------------|----------------------------------|-----------------------------------------------------------------|----------------------------------|-----------------------------------------------------------------|-------------------------------------|
|                                 |                               | N.º de<br>extracciones/<br>total | Título<br>de log <sub>10</sub><br>medio<br>(DIE <sub>50</sub> ) | N.º de<br>extracciones/<br>total | Título<br>de log <sub>10</sub><br>medio<br>(DIE <sub>50</sub> ) |                                     |
| H5N1 wt<br>1997, 2003<br>y 2004 | 8/8                           | 8/8                              | >6,3                                                            | 8/8                              | >4,5                                                            | NA                                  |
| H5N1 ca<br>1997, 2003<br>y 2004 | 0/8                           | 0/8                              | <0,9                                                            | 0/8                              | <0,9                                                            | 0/8                                 |

Se inocularon pollos por vía intranasal con  $10^6$  DICT<sub>50</sub> de virus.



**Figura 3**

| <b>DL<sub>50</sub> en ratones</b> |                                             |
|-----------------------------------|---------------------------------------------|
| <b>A/AA/6/60 ca</b>               | <b>&gt;10<sup>7</sup> DICT<sub>50</sub></b> |
| <b>A/HK/491/97</b>                | <b>10<sup>2</sup> DICT<sub>50</sub></b>     |
| <b>1997 H5N1/AA ca</b>            | <b>&gt;10<sup>7</sup> DICT<sub>50</sub></b> |
| <b>A/HK/213/2003</b>              | <b>10<sup>6</sup> DICT<sub>50</sub></b>     |
| <b>2003 H5N1/AA ca</b>            | <b>&gt;10<sup>7</sup> DICT<sub>50</sub></b> |
| <b>A/Vietnam/1203/2004</b>        | <b>10<sup>0,4</sup> DICT<sub>50</sub></b>   |
| <b>2004 H5N1/AA ca</b>            | <b>&gt;10<sup>7</sup> DICT<sub>50</sub></b> |

Figura 4

| Tejido              | Virus     | Veces promedio<br>de diferencia en título<br>a lo largo de 3 días |
|---------------------|-----------|-------------------------------------------------------------------|
| PULMONES            | A/AA/6/60 | 93                                                                |
|                     | 1997 H5N1 | 501                                                               |
|                     | 2003 H5N1 | 12                                                                |
|                     | 2004 H5N1 | 430                                                               |
| CORNETES<br>NASALES | A/AA/6/60 | 32                                                                |
|                     | 1997 H5N1 | 185                                                               |
|                     | 2003 H5N1 | ninguna                                                           |
|                     | 2004 H5N1 | 100                                                               |

Se administró por vía intranasal  $10^6$  DICT<sub>50</sub> de virus y se recogieron los tejidos en los días 2, 3 ó 4 tras la infección. Los títulos de virus se expresan como log<sub>10</sub> DICT<sub>50</sub>/g de tejido.

**Figura 5**  
**Títulos de virus en pulmones**

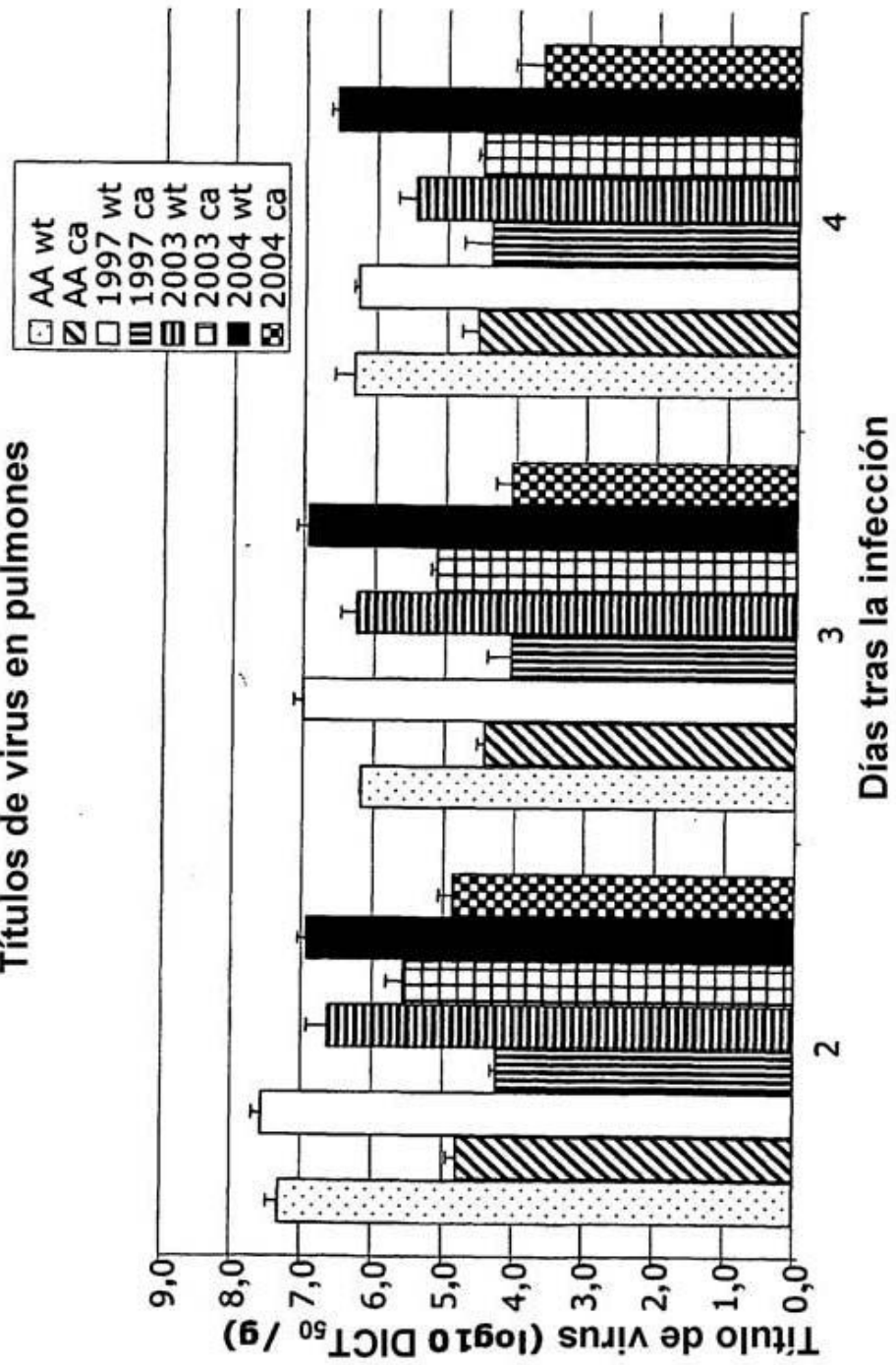


Figura 6

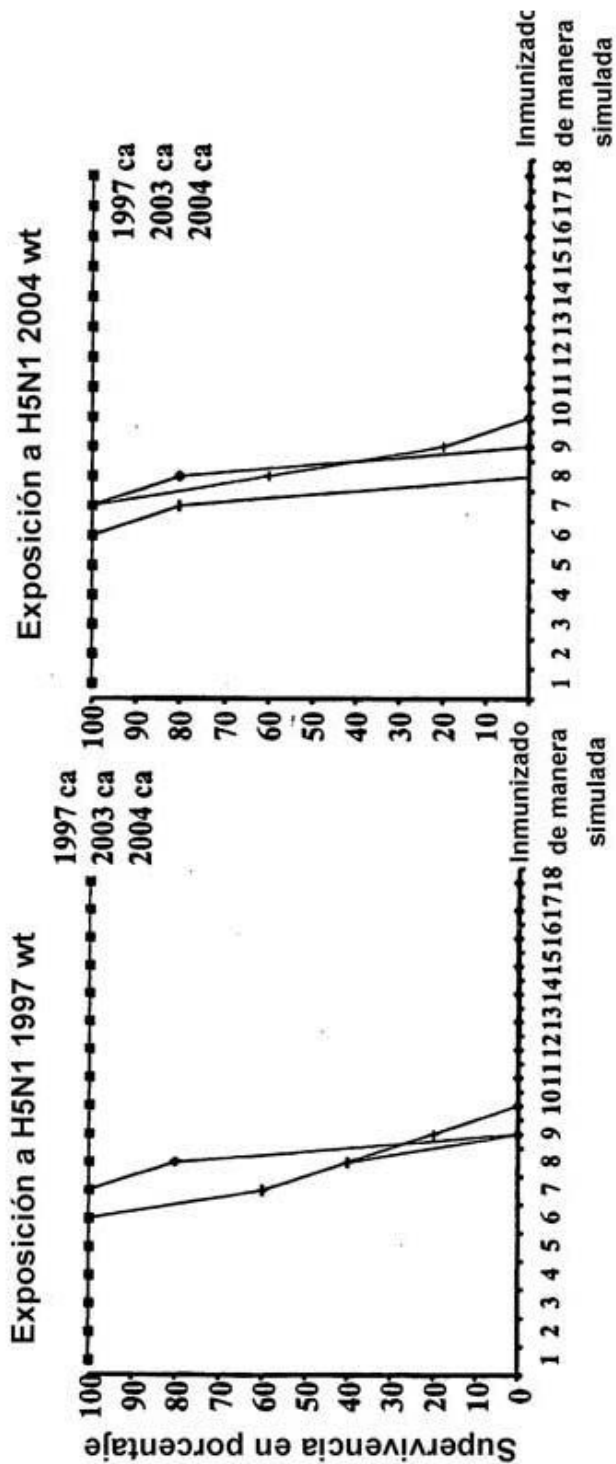
| Virus de<br>inmunización | Media geométrica de títulos de Ac HAI<br>séricos contra los virus indicados |         |         |
|--------------------------|-----------------------------------------------------------------------------|---------|---------|
|                          | 1997 wt                                                                     | 2003 wt | 2004 wt |
| 2003 ca                  | 20                                                                          | 213,6   | 20      |
| 2003 wt                  | 20                                                                          | 394     | 20      |

A un título indetectable se le asigna un valor de 20

**Figura 7**

| <b>Virus de<br/>inmunización</b>                            | <b>Media geométrica de títulos de Ac neutralizantes<br/>séricos contra los virus indicados</b> |                |                |
|-------------------------------------------------------------|------------------------------------------------------------------------------------------------|----------------|----------------|
|                                                             | <b>1997 wt</b>                                                                                 | <b>2003 wt</b> | <b>2004 wt</b> |
| <b>2003 ca</b>                                              | <b>10</b>                                                                                      | <b>59,2</b>    | <b>10</b>      |
| <b>2003 wt</b>                                              | <b>10</b>                                                                                      | <b>93,3</b>    | <b>10</b>      |
| <b>A un título indetectable se le asigna un valor de 10</b> |                                                                                                |                |                |

Figura 8



Días tras la administración del virus de exposición

Figura 9

| Reducción media en el título en pulmones tras |                              |                               |         |         |
|-----------------------------------------------|------------------------------|-------------------------------|---------|---------|
| Inmunización                                  | Exposición<br>a<br>homólogos | Exposición a H5N1 heterólogos |         |         |
|                                               |                              | 1997 wt                       | 2003 wt | 2004 wt |
| 1997 ca                                       | 2,5                          | NA                            | 3,0     | 0,7     |
| 2003 ca                                       | >5,8                         | 2,3                           | NA      | 2,9     |
| 2004 ca                                       | 2,0                          | 1,4                           | >5,7    | NA      |

Figura 10

| Reducción media en el título en CN tras |                                      |                               |         |         |
|-----------------------------------------|--------------------------------------|-------------------------------|---------|---------|
| Inmunización                            | Exposición <sup>a</sup><br>homólogos | Exposición a H5N1 heterólogos |         |         |
|                                         |                                      | 1997 wt                       | 2003 wt | 2004 wt |
| 1997 ca                                 | 4,3                                  | NA                            | >1,2    | 2,6     |
| 2003 ca                                 | >1,2                                 | 3,7                           | NA      | >3,3    |
| 2004 ca                                 | 1,6                                  | 4,2                           | >3,5    | NA      |



Figura 11

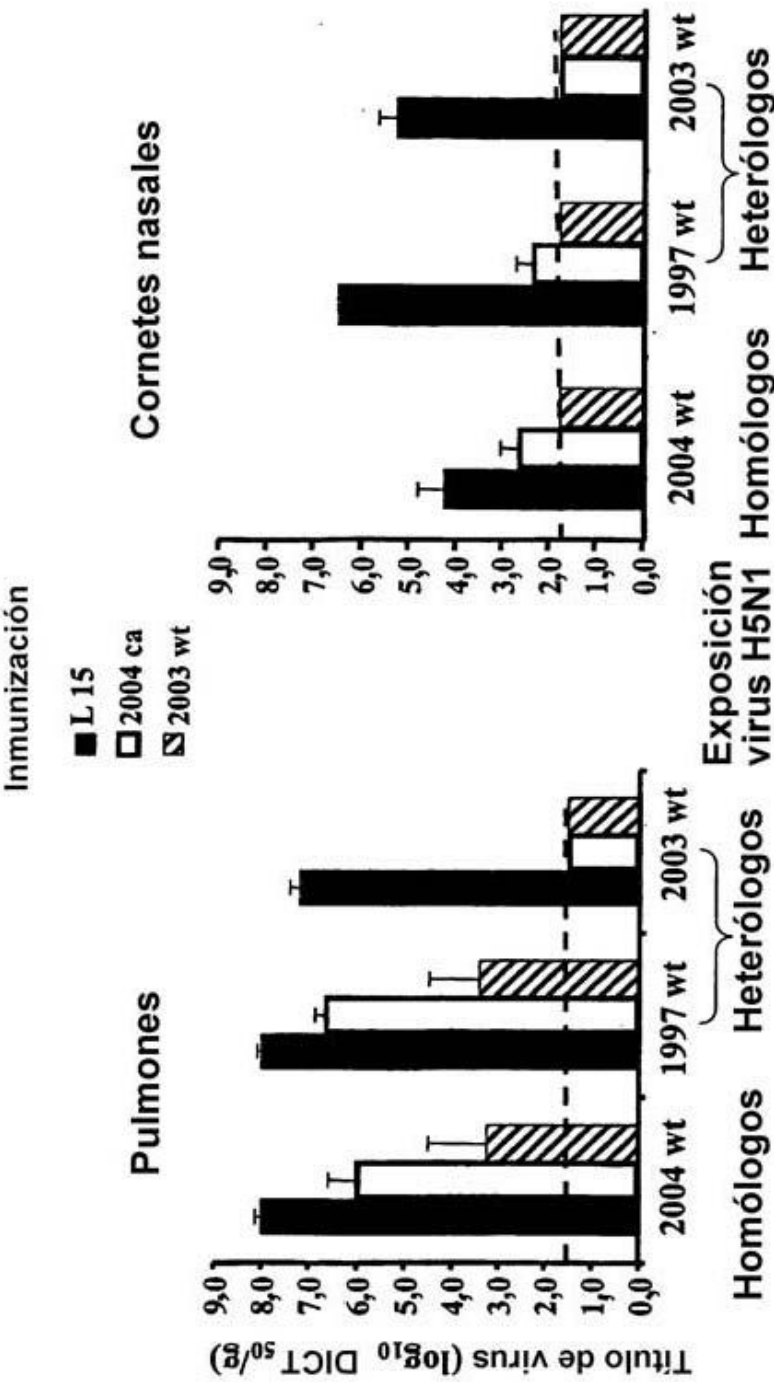


Figura 12

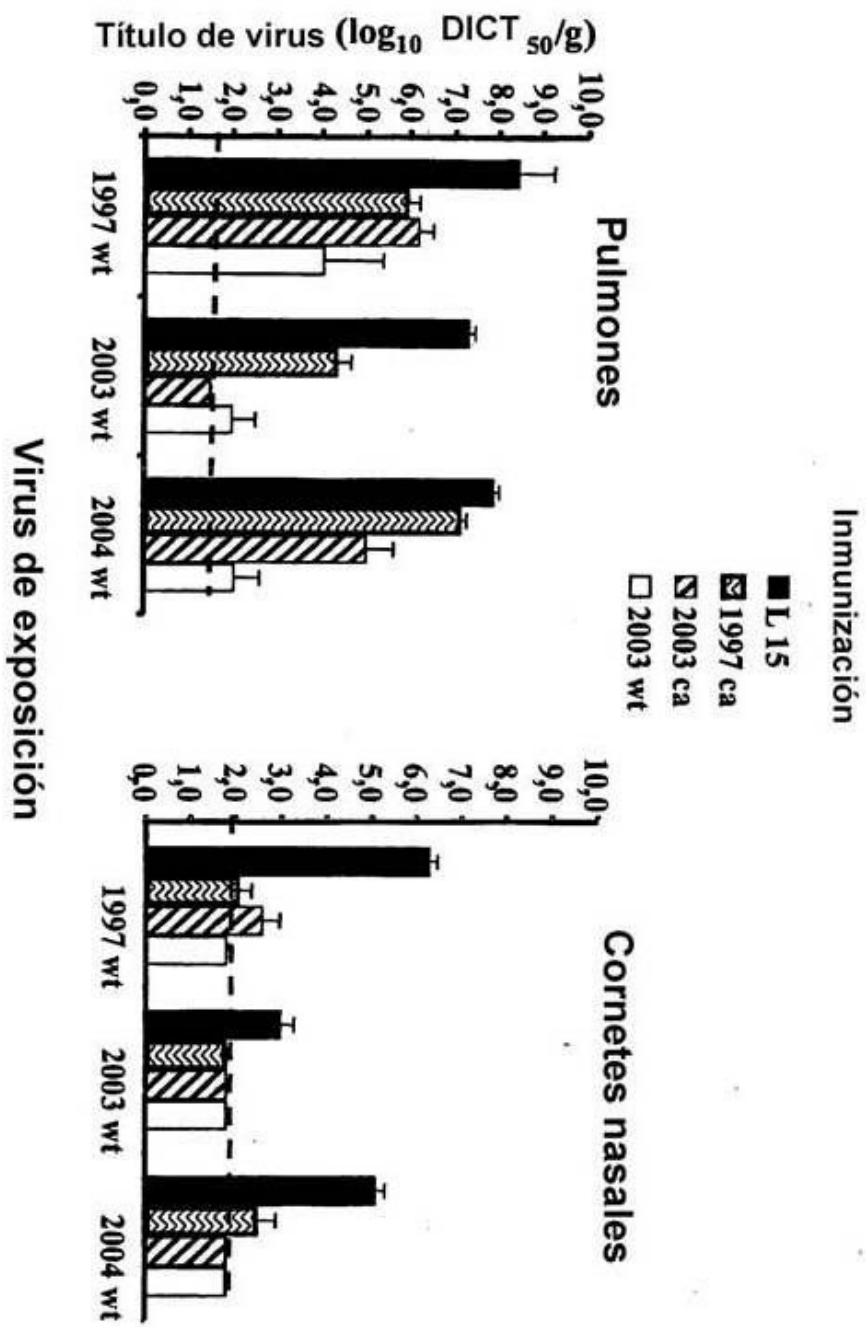


Figura 13

