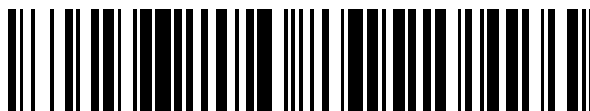


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 662 030**

51 Int. Cl.:

C12N 15/11

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **02.10.2008** **PCT/AU2008/001465**

87 Fecha y número de publicación internacional: **23.04.2009** **WO09049351**

96 Fecha de presentación y número de la solicitud europea: **02.10.2008** **E 08800100 (3)**

97 Fecha y número de publicación de la concesión europea: **13.12.2017** **EP 2215269**

54 Título: **Sistema de construcciones y usos para el mismo**

30 Prioridad:

15.10.2007 US 980145 P

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

05.04.2018

73 Titular/es:

ADMEDUS VACCINES PTY LIMITED (100.0%)
PO Box 836
Stones Corner, QLD 4102, AU

72 Inventor/es:

FRAZER, IAN HECTOR y
DUTTON, JULIE

74 Agente/Representante:

ISERN JARA, Jorge

ES 2 662 030 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

Sistema de construcciones y usos para el mismo

5 Campo de la invención

La presente invención se refiere en general a la expresión génica. Más en particular, la presente invención se refiere a sistemas de construcciones y a métodos para la comparación de diferentes codones iso-aceptores de acuerdo con su preferencia para la traducción de transcritos de ARN en proteínas en células o tejidos de interés o para la producción de un fenotipo seleccionado en un organismo de interés o parte del mismo. Las comparaciones de la preferencia de codón obtenidas de este modo son particularmente útiles para la modificación de la eficiencia de traducción de polinucleótidos que codifican proteínas en células o tejidos de interés o para la modulación de la cualidad de un fenotipo seleccionado conferido por un polipéptido asociado a fenotipo en un organismo de interés o parte del mismo.

15 Antecedentes de la invención

La expresión de genes heterólogos extraños en células transformadas es ahora cosa frecuente. Un gran número de genes de mamíferos, incluyendo, por ejemplo, genes murinos y humanos, se han expresado satisfactoriamente en diversas células hospedadoras, incluyendo células hospedadoras bacterianas, de levaduras, de insectos, de plantas y de mamíferos. Sin embargo, a pesar del conocimiento creciente de sistemas de expresión y de la tecnología del ADN recombinante, siguen existiendo obstáculos significativos cuando se intenta expresar un gen extraño o sintético en una célula hospedadora seleccionada. Por ejemplo, la traducción de un gen sintético, incluso cuando se acopla con un promotor fuerte, con frecuencia transcurre mucho más lentamente de lo esperado. Lo mismo es cierto frecuentemente con respecto a genes exógenos que son ajenos a la célula hospedadora. Esta eficiencia de traducción menor de la esperada es con frecuencia debido a las regiones codificantes de proteínas de los genes que tienen un patrón de uso de codones que no se asemeja al de genes altamente expresados en la célula hospedadora. Se sabe a este respecto que la utilización de los codones está altamente sesgada y varía considerablemente en diferentes organismos y que los sesgos en el uso de codones pueden alterar la tasa de elongación de péptidos. También se sabe que los patrones de uso de codones se relacionan con la abundancia relativa de isoaceptores de ARNt y que los genes que codifican proteínas de abundancia alta frente a baja muestran diferencias en sus preferencias de codón.

Las implicaciones de los fenómenos de preferencia de codón en la expresión génica se manifiestan en que estos fenómenos pueden afectar a la eficiencia de traducción del ARN mensajero (ARNm). Se sabe ampliamente a este respecto que la traducción de "codones raros", para los que el correspondiente iso-ARNt está en abundancia baja con relación a otros iso-ARNt, puede provocar que un ribosoma haga una pausa durante la traducción lo que puede conducir al fracaso para completar una cadena de polipéptidos naciente y a un desacoplamiento de la transcripción y la traducción. Por tanto, la expresión de un gen exógeno puede dificultarse gravemente si una célula hospedadora particular de un organismo, o el organismo en sí, tiene una abundancia baja de iso-ARNt que correspondes a uno o más codones del gen exógeno. En consecuencia, un objetivo primordial de los investigadores en este campo es determinar, en primer lugar, la preferencia de codón para células particulares en las que se ha de expresar un gen exógeno y, posteriormente, alterar la composición de codones de ese gen para la expresión optimizada en esas células.

Las técnicas de optimización de codones son conocidas por mejorar la cinética de traducción de regiones codificantes de proteínas traducionalmente ineficientes. Tradicionalmente, estas técnicas se han basado en el reemplazo de codones que se usan rara vez o con poca frecuencia en la célula hospedadora por aquellos preferidos por el hospedador. Las frecuencias de los codones pueden derivar de fuentes bibliográficas para los genes altamente expresados de muchos organismos (véase, por ejemplo, Nakamura et al., 1996, *Nucleic Acids Res* 24: 214-215). Estas frecuencias se expresan en general 'basándose en el promedio de todo el organismo' como el porcentaje de ocasiones en que se usa un codón sinónimo para codificar un aminoácido correspondiente en una colección de genes que codifican proteínas de dicho organismo, que preferentemente se expresan altamente.

Normalmente, los codones se clasifican como: (a) codones "comunes" (o codones "preferidos") si su frecuencia de uso está por encima de aproximadamente $4/3$ x la frecuencia de uso que se esperaría en ausencia de cualquier sesgo en el uso de los codones; (b) codones "raros" (o codones "no preferidos") si su frecuencia de uso está por debajo de aproximadamente $2/3$ x la frecuencia de uso que se esperaría en ausencia de cualquier sesgo en el uso de los codones; y (c) codones "intermedios" (o "codones menos preferidos") si su frecuencia de uso está entre la frecuencia de uso de los codones "comunes" y la de los codones "raros". Puesto que un aminoácido puede estar codificado por 2, 3, 4 o 6 codones, la frecuencia de uso de cualquier codón seleccionado, que se esperaría en ausencia de cualquier sesgo en el uso de codones, será dependiente de la cantidad de codones sinónimos que codifican el mismo aminoácido que el codón seleccionado. En consecuencia, para un aminoácido particular, los umbrales de frecuencia para clasificar codones en las categorías "comunes", "intermedios" y "raros" dependen del número de codones sinónimos para ese aminoácido. Como consecuencia, para aminoácidos que tienen 6 elecciones de codones sinónimos, la frecuencia de uso de codón que se esperaría en ausencia de cualquier sesgo

en el uso de codón es del 16 % y, por tanto, los codones "comunes", "intermedios" y "raros" se definen como aquellos codones que tienen una frecuencia de uso por encima del 20 %, entre el 10 % y el 20 % y por debajo del 10 %, respectivamente. Para los aminoácidos que tienen 4 elecciones de codones sinónimos, la frecuencia de uso de codón que se esperaría en ausencia de sesgo del uso de codón es del 25 % y, por tanto, los codones "comunes", "intermedios" y "raros" se definen como aquellos codones que tienen una frecuencia de uso por encima del 33 %, entre el 16 y el 33 % y por debajo del 16 %, respectivamente. Para la isoleucina, que es el único aminoácido que tiene 3 elecciones de codones sinónimos, la frecuencia de uso de codón que se esperaría en ausencia de cualquier sesgo en el uso de codón es del 33 % y, por tanto, los codones "comunes", "intermedios" y "raros" se definen como aquellos codones que tienen una frecuencia de uso por encima del 45 %, entre el 20 y el 45 % y por debajo del 20 %, respectivamente. Para los aminoácidos que tienen 2 elecciones de codones sinónimos, la frecuencia de uso de codón que se esperaría en ausencia de sesgo del uso de codón es del 50 % y, por tanto, los codones "comunes", "intermedios" y "raros" se definen como aquellos codones que tienen una frecuencia de uso por encima del 60 %, entre el 30 y el 60 % y por debajo del 30 %, respectivamente. Por tanto, la categorización de codones en las clases "comunes", "intermedios" y "raros" (o "preferidos", "menos preferidos" o "no preferidos", respectivamente) se han basado convencionalmente en una compilación de uso de codones para un organismo en general (por ejemplo, 'seres humanos') o para una clase de organismos en general (por ejemplo, 'mamíferos'). Por ejemplo, puede hacerse referencia a Seed (véanse las Patentes de los EE.UU. N.º de Serie 5.786.464 y 5.795.737) que desvela codones preferidos, menos preferidos y no preferidos para células de mamíferos en general. Sin embargo, el presente inventor reveló en el documento WO 99/02694 y en el documento WO 00/42190 que existen diferencias sustanciales en la abundancia relativa de iso-ARNt particulares en diferentes células o tejidos de un único organismo multicelular (por ejemplo, un mamífero o una planta) y que esto desempeña una función fundamental en la traducción de proteínas a partir de una secuencia codificante con un uso o composición de codones dados.

De este modo, de forma contraria a la suposición reconocida en la técnica de que células diferentes de un organismo multicelular tienen el mismo sesgo en el uso de codones, se reveló por primera vez que un solo tipo celular de un organismo multicelular usa codones de una manera distinta que otro tipo celular del mismo organismo. En otras palabras, se descubrió que las diferentes células de un organismo pueden presentar eficiencias de traducción diferentes para el mismo codón y que no era posible predecir que codones serían preferidos, menos preferidos o no preferidos en un tipo celular seleccionado. En consecuencia, se propuso que las diferencias en la eficiencia de traducción de codones entre los tipos celulares podrían aprovecharse, junto con la composición de codones de un gen, para regular la producción de una proteína en, o para dirigir esa producción a, un tipo celular elegido.

Por tanto, con el fin de optimizar la expresión de un polinucleótido que codifica proteína en un tipo celular particular, el documento WO 99/02694 y el documento WO 00/42190 enseñan que es necesario determinar, en primer lugar, la eficiencia de traducción para cada codón en ese tipo celular, en lugar de depender de las frecuencias de codones calculadas basándose en el promedio de todo el organismo y, después, modificar los codones del polinucleótido basándose en dicha determinación. El documento WO 00/42190 enseña adicionalmente un sistema de vectores para la clasificación de codones sinónimos de acuerdo con sus eficiencias de traducción. Este sistema de vectores comprende una pluralidad de construcciones sintéticas, comprendiendo, cada una, una secuencia reguladora que se une operativamente a una repetición en tándem de un codón fusionado en fase con un polinucleótido indicador que codifica una proteína indicadora, en el que el codón repetido en tándem de una construcción es diferente al codón repetido en tándem de otra. En este sistema, se cree que el codón repetido en tándem provoca que un ribosoma haga una pausa durante la traducción si el iso-ARNt que corresponde al codón repetido en tándem es limitante. En consecuencia, los niveles de proteína indicadora producida usando este sistema de vectores son sensibles a la abundancia intracelular de las especies de iso-ARNt que corresponden al codón repetido en tándem y proporcionan, por tanto, una correlación directa de la preferencia de célula o tejido para traducir un codón dado. Esto significa, por ejemplo, que si los niveles de la proteína indicadora obtenida en un tipo celular o tisular al que se proporciona una construcción sintética que tiene un primer codón repetido en tándem son más bajos que los niveles expresados en el mismo tipo celular o tisular al que se proporciona una construcción sintética diferente que tiene un segundo codón repetido en tándem (es decir, en los que el primer codón repetido en tándem es diferente, pero sinónimo, del segundo codón repetido en tándem), entonces puede deducirse que el segundo codón repetido en tándem tiene una eficiencia de traducción mayor que el primer codón repetido en tándem en el tipo celular o tisular.

El presente inventor determinó, adicionalmente, una estrategia para potenciar o reducir la cualidad de un fenotipo seleccionado (inmunidad, tolerancia, resistencia a patógenos, potenciación o prevención de un proceso de reparación, resistencia a las plagas, resistencia a las heladas, tolerancia a herbicidas, etc.) presentado, o que se propone que sea presentado, por un organismo de interés. Esta estrategia, que se desvela en el documento WO 2004/042059, implica la modificación de codones de un polinucleótido que codifica un polipéptido asociado a fenotipo que en el organismo de interés ya sea por sí mismo, o en asociación con otras moléculas, transmite o confiere el fenotipo seleccionado al organismo. A diferencia de métodos anteriores, que dependen de datos que proporcionen una clasificación de codones sinónimos de acuerdo con su preferencia de uso o de acuerdo con sus eficiencias de traducción, esta estrategia se basa en la clasificación de codones sinónimos individuales de acuerdo con su preferencia de uso por el organismo o clase de organismos, o por una parte de los mismos, para producir el fenotipo seleccionado. Un método ilustrativo para la determinación de preferencias fenotípicas de codones se desvela en el documento WO 2004/042059, que emplea el sistema de construcciones sintético desvelado en el

documento WO 00/42190 para derivar un conjunto de codones sinónimos que pueda mostrar una gama de preferencias fenotípicas, que puede usarse como una base para seleccionar racionalmente un codón en un polinucleótido que codifica un polipéptido asociado a fenotipo para el reemplazo con un codón sinónimo que tiene una preferencia fenotípica diferente.

Sumario de la invención

La presente invención se basa en parte en el descubrimiento de que la sensibilidad de la determinación de la eficiencia de traducción o preferencia fenotípica de diferentes codones sinónimos puede mejorarse usando un sistema de construcciones que emplea diferentes polinucleótidos indicadores que codifican la misma secuencia de aminoácidos, en el que polinucleótidos indicadores individuales usan el mismo codón (también denominado en el presente documento "un codón de interrogación") para codificar un aminoácido particular en una o más posiciones de la secuencia de aminoácidos y en el que el codón de interrogación de un polinucleótido indicador es diferente, pero sinónimo, del codón de interrogación de otro polinucleótido indicador. En realizaciones específicas, la sensibilidad mejora adicionalmente mediante la incorporación de dos o más codón de interrogación para codificar el aminoácido particular en la secuencia de aminoácidos.

Por tanto, en un aspecto de la presente invención, se proporcionan sistemas de construcciones para determinar la eficiencia de traducción o la preferencia fenotípica de diferentes codones sinónimos. Estos sistemas en general comprenden una pluralidad de construcciones sintéticas, comprendiendo, cada una, una secuencia reguladora que está conectada operativamente a un polinucleótido indicador, en las que el polinucleótido indicador de una primera construcción comprende una primera secuencia codificante para interrogar la eficiencia de traducción o la preferencia fenotípica de un primer codón ("el primer codón de interrogación") que codifica un primer aminoácido, en las que el polinucleótido indicador de una segunda construcción comprende una segunda secuencia codificadora para interrogar la eficiencia de traducción o la preferencia fenotípica de un segundo codón ("el segundo codón de interrogación") que codifica el primer aminoácido, en las que las secuencias codificantes primera y segunda codifican la misma secuencia de aminoácidos, en las que la primera secuencia codificante comprende el primer codón de interrogación para codificar el primer aminoácido en una o más posiciones de la secuencia de aminoácidos, en las que la segunda secuencia codificante comprende el segundo codón de interrogación para codificar el primer aminoácido en una o más posiciones de la secuencia de aminoácidos y en las que las secuencias de codificación primera y segunda difieren entre sí en la elección del primer codón de interrogación o el segundo codón de interrogación para codificar el primer aminoácido en la posición o posiciones correspondientes en la secuencia de aminoácidos. Convenientemente, la primera secuencia codificante comprende el mismo número de primeros codones de interrogación que el número de segundos codones de interrogación en la segunda secuencia codificante. En realizaciones específicas, la segunda secuencia codificante difiere de la primera segunda secuencia codificante en el reemplazo del primer codón de interrogación por el segundo codón de interrogación para codificar el primer aminoácido en la una o más posiciones de la secuencia de aminoácidos. En algunas realizaciones, el sistema de construcciones comprende una o más construcciones sintéticas para interrogar la eficiencia de traducción o la preferencia fenotípica de uno o más codones de interrogación que codifican el primer aminoácido. En ejemplos ilustrativos de este tipo, el sistema de construcciones comprende un número correspondiente de construcciones sintéticas como el número de codones sinónimos que normalmente codifican el primer aminoácido.

En algunas realizaciones, la secuencia codificante de construcciones sintéticas individuales comprende al menos 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400, 500 codones de interrogación del tipo correspondiente. Convenientemente, al menos el 10 %, el 15 %, el 20 %, el 25 %, el 30 %, el 35 %, el 40 %, el 45 %, el 50 %, el 55 %, el 60 %, el 65 %, el 70 %, el 75 %, el 80 %, el 85 %, el 90 %, el 95 %, el 99 % o incluso el 100 % de los codones que codifican el primer aminoácido en la secuencia codificante de construcciones sintéticas individuales son el mismo codón de interrogación.

En algunas realizaciones, el sistema de construcciones comprende adicionalmente una tercera construcción y una cuarta construcción, en las que el polinucleótido indicador de la tercera construcción comprende una tercera secuencia codificante para interrogar la eficiencia de traducción o la preferencia fenotípica de un tercer codón ("el tercer codón de interrogación") que codifica un segundo aminoácido que es diferente del primer aminoácido, en las que el polinucleótido indicador de la cuarta construcción comprende una cuarta secuencia codificante para interrogar la eficiencia de traducción o la preferencia fenotípica de un cuarto codón ("el cuarto codón de interrogación") que codifica el segundo aminoácido, en las que las secuencias codificantes tercera y cuarta codifican la misma secuencia de aminoácidos que las secuencias de codificación primera y segunda, en las que la tercera secuencia codificante comprende el tercer codón de interrogación para codificar el segundo aminoácido en una o más posiciones de la secuencia de aminoácidos, en las que la cuarta secuencia codificante comprende el cuarto codón de interrogación para codificar el segundo aminoácido en una o más posiciones de la secuencia de aminoácidos y en las que las secuencias de codificación tercera y cuarta difieren entre sí en la elección del tercer codón de interrogación o el cuarto codón de interrogación para codificar el segundo aminoácido en la posición o posiciones correspondientes en la secuencia de aminoácidos.

En algunas realizaciones, el sistema de construcciones comprende adicionalmente construcciones sintéticas para interrogar la eficiencia de traducción o la preferencia fenotípica de codones que codifican otros aminoácidos.

En algunas realizaciones, la secuencia codificante de polinucleótidos indicadores individuales codifica una secuencia de aminoácidos que confiere un fenotipo a una célula o tejido en los que se expresa la secuencia codificante (por ejemplo, una secuencia de aminoácidos de una proteína indicadora que, cuando está presente en una célula o tejido, es detectable ya sea por su presencia o actividad, incluyendo, pero no limitada a, una proteína indicadora quimioluminiscente tal como luciferasa, una proteína fluorescente tal como proteína fluorescente verde, una proteína indicadora enzimática tal como cloranfenicol acetil transferasa, β -galactosidasa, fosfatasa alcalina placentaria secretada, β -lactamasa o un factor de crecimiento tal como la hormona de crecimiento humano). Dichas proteínas indicadoras son útiles, por ejemplo, en la determinación de la eficiencia de traducción de diferentes codones sinónimos en un tipo celular o tisular de interés. En otras realizaciones, la secuencia codificante de polinucleótidos indicadores individuales codifica una secuencia de aminoácidos que confiere un fenotipo a una célula o tejido en los que la secuencia codificante no se expresa incluyendo, por ejemplo, la secuencia de aminoácidos de un polipéptido asociado a fenotipo que es el objeto de la producción de un fenotipo seleccionado (por ejemplo, inmunidad celular a melanoma) o un fenotipo de la misma clase que el fenotipo seleccionado (por ejemplo, una respuesta inmunitaria celular); como por ejemplo se desvela en el documento WO 2004/042059.

En algunas realizaciones, el polinucleótido indicador de construcciones sintéticas individuales comprende adicionalmente una secuencia auxiliar de codificación que codifica un marcador detectable, que es convenientemente un miembro de un par de unión específico, que incluye, por ejemplo, pares anticuerpo-antígeno (o hapteno), pares ligando-receptor, pares enzima-sustrato, pares biotina-avidina y similares. En los ejemplos ilustrativos de este tipo, la secuencia codificante auxiliar de un solo polinucleótido indicador codifica un marcador diferente de la secuencia codificante auxiliar de otro polinucleótido indicador. En estos ejemplos, es posible distinguir de forma detectable los productos polipeptídicos de diferentes polinucleótidos indicadores en la misma célula u organismo de interés o parte de los mismos, permitiendo de este modo la determinación simultánea de las eficiencias de traducción de codones de interrogación diferentes en la misma célula u organismo o parte de los mismos.

En otro aspecto, la presente divulgación proporciona métodos para determinar la eficiencia de traducción de un primer codón con respecto a un segundo codón en una célula de interés, en los que el primer codón y el segundo codón codifican el mismo aminoácido. Estos métodos en general comprenden:

- proporcionar una pluralidad de construcciones sintéticas, comprendiendo, cada una, una secuencia reguladora que está conectada operativamente a un polinucleótido indicador, en las que el polinucleótido indicador de una primera construcción comprende una primera secuencia codificante para interrogar la eficiencia de traducción del primer codón, en las que el polinucleótido indicador de una segunda construcción comprende una segunda secuencia codificante para interrogar la eficiencia de traducción del segundo codón, en las que las secuencias codificantes primera y segunda codifican la misma secuencia de aminoácidos, que define, en su totalidad o en parte, una proteína indicadora, en las que la primera secuencia codificante comprende el primer codón para codificar el primer aminoácido en una o más posiciones de la secuencia de aminoácidos, en las que la segunda secuencia codificante comprende el segundo codón para codificar el primer aminoácido en una o más posiciones de la secuencia de aminoácidos y en las que las secuencias de codificación primera y segunda difieren entre sí en la elección del primer codón o el segundo codón para codificar el primer aminoácido en la posición o posiciones correspondientes en la secuencia de aminoácidos;
- introducir la primera construcción en una célula del mismo tipo que la célula de interés;
- introducir la segunda construcción en una célula del mismo tipo que la célula de interés;
- medir la expresión de la proteína indicadora de la primera construcción y de la segunda construcción en la célula; y
- determinar la eficiencia de traducción del primer codón y la eficiencia de traducción del segundo codón basándose en la expresión medida de la proteína indicadora en la célula, para determinar de este modo la eficiencia de traducción del primer codón con respecto al segundo codón en la célula de interés.

En algunas realizaciones, las construcciones primera y segunda se introducen por separado en células diferentes. En otras realizaciones, las construcciones primera y segunda se introducen en la misma célula.

Convenientemente, los métodos comprenden adicionalmente determinar una comparación de eficiencias de traducción de codones sinónimos individuales en la célula de interés.

En algunas realizaciones, los métodos comprenden adicionalmente:

- introducir una construcción sintética individual en un progenitor de la célula de interés; y
- diferenciar la célula de interés del progenitor,

en los que la célula de interés contiene la construcción sintética.

En otro aspecto más, la presente divulgación proporciona métodos para determinar la eficiencia de traducción de un primer codón y un segundo codón en un primer tipo celular con respecto a un segundo tipo celular. Estos métodos en general comprenden:

- proporcionar una pluralidad de construcciones sintéticas, comprendiendo, cada una, una secuencia reguladora que está conectada operativamente a un polinucleótido indicador, en las que el polinucleótido indicador de una primera construcción comprende una primera secuencia codificante para interrogar la eficiencia de traducción del primer codón, en las que el polinucleótido indicador de una segunda construcción comprende una segunda secuencia codificante para interrogar la eficiencia de traducción del segundo codón, en las que las secuencias codificantes primera y segunda codifican la misma secuencia de aminoácidos, que define, en su totalidad o en parte, una proteína indicadora, en las que la primera secuencia codificante comprende el primer codón para codificar el primer aminoácido en una o más posiciones de la secuencia de aminoácidos, en las que la segunda secuencia codificante comprende el segundo codón para codificar el primer aminoácido en una o más posiciones de la secuencia de aminoácidos y en las que las secuencias de codificación primera y segunda difieren entre sí en la elección del primer codón o el segundo codón para codificar el primer aminoácido en la posición o posiciones correspondientes en la secuencia de aminoácidos;
 - introducir por separado la primera construcción en el primer tipo celular y en el segundo tipo celular;
 - introducir por separado la segunda construcción en el primer tipo celular y en el segundo tipo celular;
 - medir la expresión de la proteína indicadora en el primer tipo celular y en el segundo tipo celular a los que se proporcionó la primera construcción;
 - medir la expresión de la proteína indicadora en el primer tipo celular y en el segundo tipo celular a los que se proporcionó la segunda construcción;
 - determinar la eficiencia de traducción del primer codón en el primer tipo celular y en el segundo tipo celular basándose en la expresión medida de la proteína indicadora en el primer tipo celular y en el segundo tipo celular, respectivamente, a los que se proporcionó la primera construcción, para determinar de este modo la eficiencia de traducción del primer codón en el primer tipo celular con respecto al segundo tipo celular; y
 - determinar la eficiencia de traducción del segundo codón en el primer tipo celular y en el segundo tipo celular basándose en la expresión medida de la proteína indicadora en el primer tipo celular y en el segundo tipo celular, respectivamente, a los que se proporcionó la segunda construcción, para determinar de este modo la eficiencia de traducción del segundo codón en el primer tipo celular con respecto al segundo tipo celular.
- En algunas realizaciones, los métodos comprenden adicionalmente determinar una comparación de eficiencias de traducción de codones sinónimos individuales en el primer tipo celular con respecto al segundo tipo celular.

En algunas realizaciones, los métodos comprenden adicionalmente:

- introducir una construcción sintética individual en un progenitor de una célula seleccionada entre el primer tipo celular o el segundo tipo celular; y
- diferenciar la célula del progenitor,

en los que la célula contiene la construcción sintética.

Otro aspecto más de la presente divulgación proporciona métodos para determinar la preferencia de un primer codón con respecto a la preferencia de un segundo codón para producir un fenotipo seleccionado ("la preferencia fenotípica") en un organismo de interés o parte del mismo, en los que el primer codón y el segundo codón codifican el mismo aminoácido. Estos métodos en general comprenden:

- proporcionar una pluralidad de construcciones sintéticas, comprendiendo, cada una, una secuencia reguladora que está conectada operativamente a un polinucleótido indicador, en las que el polinucleótido indicador de una primera construcción comprende una primera secuencia codificante para interrogar la preferencia fenotípica del primer codón, en las que el polinucleótido indicador de una segunda construcción comprende una segunda secuencia codificante para interrogar la preferencia fenotípica del segundo codón, en las que las secuencias codificantes primera y segunda codifican la misma secuencia de aminoácidos, que define, en su totalidad o en parte, una proteína indicadora, que produce, o que se prevé que produce, el fenotipo seleccionado o un fenotipo de la misma clase que el fenotipo seleccionado, en las que la primera secuencia codificante comprende el primer codón para codificar el primer aminoácido en una o más posiciones de la secuencia de aminoácidos, en las que la segunda secuencia codificante comprende el segundo codón para codificar el primer aminoácido en una o más

posiciones de la secuencia de aminoácidos y en las que las secuencias de codificación primera y segunda difieren entre sí en la elección del primer codón o el segundo codón para codificar el primer aminoácido en la posición o posiciones correspondientes en la secuencia de aminoácidos;

- 5 - introducir la primera construcción en un primer organismo de ensayo o parte del mismo, en los que el organismo de ensayo se selecciona entre el grupo que consiste en un organismo de la misma especie que el organismo de interés y un organismo que está relacionado con el organismo de interés;
- 10 - introducir la segunda construcción en un segundo organismo de ensayo o parte del mismo, en los que el segundo organismo de ensayo es del mismo tipo que el primer organismo;
- determinar la cualidad del fenotipo correspondiente presentado por el primer organismo de ensayo o parte del mismo y por el segundo organismo de ensayo o parte del mismo; y
- 15 - determinar la preferencia fenotípica del primer codón y la preferencia fenotípica del segundo codón basándose, respectivamente, en la cualidad del fenotipo correspondiente presentado por el primer organismo de ensayo o parte del mismo y por el segundo organismo de ensayo o parte del mismo, para determinar de este modo la preferencia fenotípica del primer codón con respecto a la preferencia fenotípica del segundo codón en el organismo de interés o parte del mismo.

20 En algunas realizaciones, los métodos comprenden adicionalmente determinar una comparación de preferencias fenotípicas de codones sinónimos individuales en el organismo de interés o parte del mismo.

En algunas realizaciones, los métodos comprenden adicionalmente:

- 25 - introducir una construcción sintética individual en un progenitor del organismo de ensayo o parte del mismo; y
- cultivar un organismo no humano o parte del mismo a partir del progenitor,

30 en los que el organismo o parte del mismo contiene la construcción sintética.

En otras realizaciones, los métodos comprenden adicionalmente:

- 35 - introducir una construcción sintética individual en un progenitor del organismo de ensayo o parte del mismo; y
- cultivar un organismo no humano o parte del mismo a partir del progenitor,

en los que el organismo o parte del mismo comprende una célula que contiene la construcción sintética.

40 En otro aspecto más, la presente divulgación proporciona métodos de construcción de un polinucleótido sintético a partir del cual un polipéptido codificado se produce a un nivel más alto en una célula de interés que a partir de un polinucleótido parental que codifica el mismo polipéptido. Estos métodos en general comprenden:

- 45 - determinar la eficiencia de traducción de diferentes codones sinónimos en células del mismo tipo que la célula de interés, como se ha descrito ampliamente anteriormente, para determinar de este modo una comparación de eficiencias de traducción de codones sinónimos individuales en la célula de interés;
- seleccionar un primer codón del polinucleótido parental para el reemplazo con un codón sinónimo, en los que el codón sinónimo se selecciona basándose en que éste presenta una mayor eficiencia de traducción que el primer codón en la célula de interés de acuerdo con la comparación de eficiencias de traducción; y
- 50 - reemplazar el primer codón con el codón sinónimo para construir el polinucleótido sintético.

55 En algunas realizaciones, el codón sinónimo se selecciona basándose en que corresponde a un codón de interrogación en una construcción sintética a partir de la cual se expresa la proteína indicadora en la célula de interés a un nivel que es al menos de aproximadamente un 10 %, un 15 %, un 20 %, un 25 %, un 30 %, un 35 %, un 40 %, un 45 %, un 50 %, un 55 %, un 60 %, un 65 %, un 70 %, un 75 %, un 80 %, un 85 %, un 90 % o un 95 % superior o al menos aproximadamente 2, 3, 4, 5, 6, 7, 8, 9 o 10 veces mayor que el nivel de la proteína indicadora expresada a partir de una construcción sintética que comprende el primer codón como el codón de interrogación.

60 Un aspecto adicional de la presente divulgación proporciona métodos de construcción de un polinucleótido sintético a partir del cual un polipéptido codificado se produce en un nivel inferior en una célula de interés que a partir de un polinucleótido parental que codifica el mismo polipéptido. Estos métodos en general comprenden:

- determinar la eficiencia de traducción de diferentes codones sinónimos en células del mismo tipo que la célula de interés, como se ha descrito ampliamente anteriormente, para determinar de este modo una comparación de eficiencias de traducción de codones sinónimos individuales en la célula de interés;

- 5 - seleccionar un primer codón del polinucleótido parental para el reemplazo con un codón sinónimo, en los que el codón sinónimo se selecciona basándose en que éste presenta una eficiencia de traducción más baja que el primer codón en la célula de interés de acuerdo con la comparación de eficiencias de traducción; y

- reemplazar el primer codón con el codón sinónimo para construir el polinucleótido sintético.

En algunas realizaciones, el codón sinónimo se selecciona basándose en que corresponde a un codón de interrogación en una construcción sintética a partir de la cual se expresa la proteína indicadora en la célula de interés en un nivel que no es superior al 95 %, al 90 %, al 85 %, al 80 %, al 75 %, al 70 %, al 65 %, al 60 %, al 55 %, al 50 %, al 45 %, al 40 %, al 35 %, al 30 %, al 25 %, al 20 %, al 15 %, al 10 %, al 5 %, al 1 %, al 0,5 %, al 0,1 %, al 0,05 % o al 0,01 % del nivel de la proteína indicadora expresada a partir de una construcción sintética que comprende el primer codón como el codón de interrogación.

Otro aspecto más de la presente divulgación proporciona métodos de construcción de un polinucleótido sintético a partir del cual un polipéptido codificado se produce a un nivel más alto en una primera célula que en una segunda célula. Estos métodos en general comprenden:

- determinar la eficiencia de traducción de diferentes codones sinónimos en las células del mismo tipo que la primera célula y en las células del mismo tipo que la segunda célula, como se ha descrito ampliamente anteriormente, para determinar de este modo una comparación de eficiencias de traducción de codones sinónimos individuales entre la primera célula y la segunda célula;

- seleccionar un primer codón del polinucleótido progenitor para el reemplazo con un codón sinónimo, en los que el codón sinónimo se selecciona basándose en que éste presenta una mayor eficiencia de traducción en la primera célula que en la segunda célula de acuerdo con la comparación de eficiencias de traducción; y

- reemplazar el primer codón con el codón sinónimo para construir el polinucleótido sintético.

En algunas realizaciones, el codón sinónimo es el mismo que el codón de interrogación en una construcción sintética a partir de la cual la proteína indicadora se expresa en la primera célula a un nivel que es al menos aproximadamente un 10 %, un 15 %, un 20 %, un 25 %, un 30 %, un 35 %, un 40 %, un 45 %, un 50 %, un 55 %, un 60 %, un 65 %, un 70 %, un 75 %, un 80 %, un 85 %, un 90 % o un 95 % superior o al menos aproximadamente 2, 3, 4, 5, 6, 7, 8, 9 o 10 veces más alto que el nivel de la proteína indicadora expresada a partir de la misma construcción sintética en la segunda célula.

En otro aspecto más, la presente divulgación proporciona métodos de construcción de un polinucleótido sintético a partir del cual se puede producir un polipéptido para conferir un fenotipo seleccionado a un organismo de interés o parte del mismo en una cualidad diferente al conferido por un polinucleótido parental que codifica el mismo polipéptido. Estos métodos en general comprenden:

- determinar la preferencia de diferentes codones sinónimos para producir el fenotipo seleccionado ("la preferencia fenotípica") en organismos de ensayo o partes de los mismos, como se ha descrito ampliamente anteriormente, en los que los organismos de ensayo se seleccionan entre el grupo que consiste en un organismo de la misma especie que el organismo de interés y un organismo que está relacionado con el organismo de interés, para determinar de este modo una comparación de preferencias fenotípicas de codones sinónimos individuales en el organismo de interés;

- seleccionar un primer codón del polinucleótido parental para el reemplazo con un codón sinónimo, en los que el codón sinónimo se selecciona basándose en que éste presenta una preferencia fenotípica diferente a la del primer codón en la comparación de preferencias fenotípicas en el organismo o parte del mismo; y

- reemplazar el primer codón con el codón sinónimo para construir el polinucleótido sintético.

En algunas realizaciones, el polinucleótido sintético confiere el fenotipo seleccionado al organismo de interés o parte del mismo en una cualidad mayor que el conferido por el polinucleótido parental. En ejemplos ilustrativos de este tipo, el codón sinónimo se selecciona basándose en que corresponde a un codón de interrogación en una construcción sintética que confiere el fenotipo seleccionado en el organismo de interés o parte del mismo en una cualidad que es al menos aproximadamente un 10 %, un 15 %, un 20 %, un 25 %, un 30 %, un 35 %, un 40 %, un 45 %, un 50 %, un 55 %, un 60 %, un 65 %, un 70 %, un 75 %, un 80 %, un 85 %, un 90 % o un 95 % superior o al menos aproximadamente 2, 3, 4, 5, 6, 7, 8, 9 o 10 veces superior a la cualidad del fenotipo conferido por la construcción sintética que comprende el primer codón como el codón de interrogación.

En otras realizaciones, el polinucleótido sintético confiere el fenotipo seleccionado al organismo de interés o parte del mismo en una cualidad inferior que el conferido por el polinucleótido parental. En ejemplos ilustrativos de este tipo, el codón sinónimo se selecciona basándose en que corresponde a un codón de interrogación en una construcción sintética que confiere el fenotipo seleccionado en el organismo de interés o parte del mismo en una

cualidad que no es superior al 95 %, al 90 %, al 85 %, al 80 %, al 75 %, al 70 %, al 65 %, al 60 %, al 55 %, al 50 %, al 45 %, al 40 %, al 35 %, al 30 %, al 25 %, al 20 %, al 15 %, al 10 %, al 5 %, al 1 %, al 0,5 %, al 0,1 %, al 0,05 % o al 0,01 % de la cualidad del fenotipo conferido por la construcción sintética que comprende el primer codón como el codón de interrogación.

El sistema de construcciones de la presente invención se ha utilizado para determinar una clasificación de codones sinónimos individuales de acuerdo con su preferencia para la producción de una respuesta inmunitaria, incluyendo una respuesta inmunitaria humoral, a un antígeno en un mamífero. Significativamente, esta clasificación no coincide con una clasificación de valores de frecuencia de codones derivable de un análisis de la frecuencia a la que se usan los codones para codificar sus aminoácidos correspondiente en una colección de genes que codifican proteínas de mamíferos altamente expresadas, como por ejemplo se desvela por Seed (citado anteriormente). Tampoco coincide con una clasificación de valores de eficiencia de traducción obtenidos a partir de un análisis de las eficiencias de traducción de codones en tipos celulares específicos, como se desvela por ejemplo en el documento WO 99/02694 para células COS-1 y células epiteliales y en el documento WO 2004/024915 para células CHO. Como resultado, la presente invención permite, por primera vez, la construcción de polinucleótidos que codifican antígenos, que se optimizan por codones para la producción eficiente de respuestas inmunitarias, incluyendo respuestas inmunitarias humorales, en un mamífero.

En consecuencia, en otro aspecto, se proporcionan métodos para construir un polinucleótido sintético a partir del cual se puede producir un polipéptido para conferir una respuesta inmunitaria a un antígeno diana en un mamífero en una cualidad diferente a la conferida por un polinucleótido parental que codifica el mismo polipéptido, en los que el polipéptido corresponde a al menos una porción del antígeno diana. Estos métodos en general comprenden: (a) seleccionar un primer codón del polinucleótido parental para el reemplazo con un codón sinónimo, en los que el codón sinónimo se selecciona basándose en que presenta una preferencia diferente para conferir una respuesta inmunitaria ("una preferencia de respuesta inmunitaria") que el primer codón en una comparación de preferencias de respuesta inmunitaria; y (b) reemplazar el primer codón con el codón sinónimo para construir el polinucleótido sintético, en los que la comparación de las preferencias de respuesta inmunitaria de los codones se representa mediante la TABLA 1:

TABLA 1

Aminoácido	Clasificación de preferencia de respuesta inmunitaria para codones sinónimos
Ala	$\text{Ala}^{\text{GCT}} > \text{Ala}^{\text{GCC}} > (\text{Ala}^{\text{GCA}}, \text{Ala}^{\text{GCG}})$
Arg	$(\text{Arg}^{\text{CGA}}, \text{Arg}^{\text{CGC}}, \text{Arg}^{\text{CGT}}, \text{Arg}^{\text{AGA}}) > (\text{Arg}^{\text{AGG}}, \text{Arg}^{\text{CGG}})$
Asn	$\text{Asn}^{\text{AAC}} > \text{Asn}^{\text{AAT}}$
Asp	$\text{Asp}^{\text{GAC}} > \text{Asp}^{\text{GAT}}$
Cys	$\text{Cys}^{\text{TGC}} > \text{Cys}^{\text{TGT}}$
Glu	$\text{Glu}^{\text{GAA}} > \text{Glu}^{\text{GAG}}$
Gln	$\text{Gln}^{\text{CAA}} = \text{Gln}^{\text{CAG}}$
Gly	$\text{Gly}^{\text{GGA}} > (\text{Gly}^{\text{GGG}}, \text{Gly}^{\text{GGT}}, \text{Gly}^{\text{GGC}})$
His	$\text{His}^{\text{CAC}} = \text{His}^{\text{CAT}}$
Ile	$\text{Ile}^{\text{ATC}} >> \text{Ile}^{\text{ATT}} > \text{Ile}^{\text{ATA}}$
Leu	$(\text{Leu}^{\text{CTG}}, \text{Leu}^{\text{CTC}}) > (\text{Leu}^{\text{CTA}}, \text{Leu}^{\text{CTT}}) >> \text{Leu}^{\text{TTG}} > \text{Leu}^{\text{TTA}}$
Lys	$\text{Lys}^{\text{AAG}} = \text{Lys}^{\text{AAA}}$
Phe	$\text{Phe}^{\text{TTT}} > \text{Phe}^{\text{TTC}}$
Pro	$\text{Pro}^{\text{CCC}} > \text{Pro}^{\text{CCT}} >> (\text{Pro}^{\text{CCA}}, \text{Pro}^{\text{CCG}})$
Ser	$\text{Ser}^{\text{TCG}} >> (\text{Ser}^{\text{TCT}}, \text{Ser}^{\text{TCA}}, \text{Ser}^{\text{TCC}}) >> (\text{Ser}^{\text{AGC}}, \text{Ser}^{\text{AGT}})$
Thr	$\text{Thr}^{\text{ACG}} > \text{Thr}^{\text{ACC}} >> \text{Thr}^{\text{ACA}} > \text{Thr}^{\text{ACT}}$
Tyr	$\text{Tyr}^{\text{TAC}} > \text{Tyr}^{\text{TAT}}$
Val	$(\text{Val}^{\text{GTG}}, \text{Val}^{\text{GTC}}) > \text{Val}^{\text{GTT}} > \text{Val}^{\text{GTA}}$

De este modo, puede conseguirse una respuesta inmunitaria más fuerte o potenciada para el antígeno diana (por ejemplo, una respuesta inmunitaria que sea al menos aproximadamente el 110 %, el 150 %, el 200 %, el 300 %, el 400 %, el 500 %, el 600 %, el 700 %, el 800 %, el 900 %, el 1000 % y todos los porcentajes enteros entre medias, de la que se produce a partir del polinucleótido parental en condiciones idénticas) seleccionando un codón sinónimo que tenga una preferencia de respuesta inmunitaria mayor que el primer codón que reemplaza. En realizaciones

específicas, el codón sinónimo se selecciona de manera que tenga una preferencia de respuesta inmunitaria mayor que sea al menos aproximadamente un 10 % (y de al menos aproximadamente un 11 % a al menos aproximadamente un 1000 % y todos los porcentajes enteros entre medias) más alta que la preferencia de respuesta inmunitaria del codón que reemplaza. En ejemplos ilustrativos de este tipo, los codones primero y sinónimo se seleccionan de la TABLA 2:

5

TABLA 2

Primer codón	Codón sinónimo	Primer codón	Codón sinónimo	Primer codón	Codón sinónimo
Ala ^{GCG}	Ala ^{GCT}	Ile ^{ATA}	Ile ^{ATC}	Ser ^{AGT}	Ser ^{TCG}
Ala ^{GCG}	Ala ^{GCC}	Ile ^{ATA}	Ile ^{ATT}	Ser ^{AGT}	Ser ^{TCT}
Ala ^{GCA}	Ala ^{GCT}	Ile ^{ATT}	Ile ^{ATC}	Ser ^{AGT}	Ser ^{TCA}
Ala ^{GCA}	Ala ^{GCC}			Ser ^{AGT}	Ser ^{TCC}
Ala ^{GCC}	Ala ^{GCT}	Leu ^{TTA}	Leu ^{CTG}	Ser ^{AGC}	Ser ^{TCG}
		Leu ^{TTA}	Leu ^{CTC}	Ser ^{AGC}	Ser ^{TCT}
Arg ^{CGG}	Arg ^{CGA}	Leu ^{TTA}	Leu ^{CTA}	Ser ^{AGC}	Ser ^{TCA}
Arg ^{CGG}	Arg ^{CGC}	Leu ^{TTA}	Leu ^{CTT}	Ser ^{AGC}	Ser ^{TCC}
Arg ^{CGG}	Arg ^{CGT}	Leu ^{TTA}	Leu ^{TTG}	Ser ^{TCC}	Ser ^{TCG}
Arg ^{CGG}	Arg ^{AGA}	Leu ^{TTG}	Leu ^{CTG}	Ser ^{TCA}	Ser ^{TCG}
Arg ^{AGG}	Arg ^{CGA}	Leu ^{TTG}	Leu ^{CTC}	Ser ^{TCT}	Ser ^{TCG}
Arg ^{AGG}	Arg ^{CGC}	Leu ^{TTG}	Leu ^{CTA}		
Arg ^{AGG}	Arg ^{CGT}	Leu ^{TTG}	Leu ^{CTT}	Thr ^{ACT}	Thr ^{ACG}
Arg ^{AGG}	Arg ^{AGA}	Leu ^{CTT}	Leu ^{CTG}	Thr ^{ACT}	Thr ^{ACC}
		Leu ^{CTT}	Leu ^{CTC}	Thr ^{ACA}	Thr ^{ACA}
Asn ^{AAT}	Asn ^{AAC}	Leu ^{CTA}	Leu ^{CTG}	Thr ^{ACA}	Thr ^{ACG}
		Leu ^{CTA}	Leu ^{CTC}	Thr ^{ACA}	Thr ^{ACC}
Asp ^{GAT}	Asp ^{GAC}			Thr ^{ACC}	Thr ^{ACG}
		Phe ^{TTC}	Phe ^{TTT}		
Cys ^{TGT}	Cys ^{TGC}			Tyr ^{TAT}	Tyr ^{TAC}
		Pro ^{CCG}	Pro ^{CCC}		
Glu ^{GAG}	Glu ^{GAA}	Pro ^{CCG}	Pro ^{CCT}	Val ^{GTA}	Val ^{GTG}
		Pro ^{CCA}	Pro ^{CCC}	Val ^{GTA}	Val ^{GTC}
Gly ^{GGC}	Gly ^{GGA}	Pro ^{CCA}	Pro ^{CCT}	Val ^{GTA}	Val ^{GTT}
Gly ^{GGT}	Gly ^{GGA}	Pro ^{CCT}	Pro ^{CCC}	Val ^{GTT}	Val ^{GTG}
Gly ^{GGG}	Gly ^{GGA}			Val ^{GTT}	Val ^{GTC}

En otros ejemplos ilustrativos de este tipo, los codones primero y sinónimo se seleccionan de la Tabla 3:

10

TABLA 3

Primer codón	Codón sinónimo	Primer codón	Codón sinónimo	Primer codón	Codón sinónimo
Ala ^{GCG}	Ala ^{GCT}	Leu ^{TTA}	Leu ^{CTA}	Ser ^{AGT}	Ser ^{TCG}
Ala ^{GCA}	Ala ^{GCT}	Leu ^{TTA}	Leu ^{CTT}	Ser ^{AGT}	Ser ^{TCT}
Ala ^{GCC}	Ala ^{GCT}	Leu ^{TTA}	Leu ^{TTG}	Ser ^{AGT}	Ser ^{TCA}
		Leu ^{TTG}	Leu ^{CTA}	Ser ^{AGC}	Ser ^{TCG}
Arg ^{CGG}	Arg ^{CGA}	Leu ^{TTG}	Leu ^{CTT}	Ser ^{AGC}	Ser ^{TCT}
Arg ^{CGG}	Arg ^{CGT}			Ser ^{AGC}	Ser ^{TCA}
Arg ^{CGG}	Arg ^{AGA}	Phe ^{TTC}	Phe ^{TTT}	Ser ^{AGC}	Ser ^{TCC}
Arg ^{AGG}	Arg ^{CGA}			Ser ^{TCC}	Ser ^{TCG}
Arg ^{AGG}	Arg ^{CGT}	Pro ^{CCG}	Pro ^{CCT}	Ser ^{TCA}	Ser ^{TCG}
Arg ^{AGG}	Arg ^{AGA}	Pro ^{CCA}	Pro ^{CCT}	Ser ^{TCT}	Ser ^{TCG}
Glu ^{GAG}	Glu ^{GAA}			Thr ^{ACT}	Thr ^{ACG}
				Thr ^{ACT}	Thr ^{ACA}

Primer codón	Codón sinónimo	Primer codón	Codón sinónimo	Primer codón	Codón sinónimo
Gly ^{GGC}	Gly ^{GGA}			Thr ^{ACA}	Thr ^{ACG}
Gly ^{GGT}	Gly ^{GGA}			Thr ^{ACC}	Thr ^{ACG}
Gly ^{GGG}	Gly ^{GGA}				
				Val ^{GTA}	Val ^{GTT}

Convenientemente, en algunos de los ejemplos ilustrativos señalados anteriormente, el método comprende adicionalmente seleccionar un segundo codón del polinucleótido parental para el reemplazo con un codón sinónimo, en el que el codón sinónimo se selecciona basándose en que presenta una preferencia de respuesta inmunitaria mayor que el segundo codón en una comparación de preferencias de respuesta inmunitaria; y (b) reemplazar el segundo codón con el codón sinónimo, en el que la comparación de preferencias de respuesta inmunitaria de los codones se representa mediante la TABLA 4:

TABLA 4

Segundo codón	Codón sinónimo	Segundo codón	Codón sinónimo	Segundo codón	Codón sinónimo
Ala ^{GCG}	Ala ^{GCT}	Ile ^{ATA}	Ile ^{ATC}	Ser ^{AGT}	Ser ^{TCG}
Ala ^{GCG}	Ala ^{GCC}	Ile ^{ATA}	Ile ^{ATT}	Ser ^{AGT}	Ser ^{TCT}
Ala ^{GCA}	Ala ^{GCT}	Ile ^{ATT}	Ile ^{ATC}	Ser ^{AGT}	Ser ^{TCA}
Ala ^{GCA}	Ala ^{GCC}			Ser ^{AGT}	Ser ^{TCC}
Ala ^{GCC}	Ala ^{GCT}	Leu ^{TTA}	Leu ^{CTG}	Ser ^{AGC}	Ser ^{TCG}
		Leu ^{TTA}	Leu ^{CTC}	Ser ^{AGC}	Ser ^{TCT}
Arg ^{CGG}	Arg ^{CGA}	Leu ^{TTA}	Leu ^{CTA}	Ser ^{AGC}	Ser ^{TCA}
Arg ^{CGG}	Arg ^{CGC}	Leu ^{TTA}	Leu ^{CTT}	Ser ^{AGC}	Ser ^{TCC}
Arg ^{CGG}	Arg ^{CGT}	Leu ^{TTA}	Leu ^{TTG}	Ser ^{TCC}	Ser ^{TCG}
Arg ^{CGG}	Arg ^{AGA}	Leu ^{TTG}	Leu ^{CTG}	Ser ^{TCA}	Ser ^{TCG}
Arg ^{AGG}	Arg ^{CGA}	Leu ^{TTG}	Leu ^{CTC}	Ser ^{TCT}	Ser ^{TCG}
Arg ^{AGG}	Arg ^{CGC}	Leu ^{TTG}	Leu ^{CTA}		
Arg ^{AGG}	Arg ^{CGT}	Leu ^{TTG}	Leu ^{CTT}	Thr ^{ACT}	Thr ^{ACG}
Arg ^{AGG}	Arg ^{AGA}	Leu ^{CTT}	Leu ^{CTG}	Thr ^{ACT}	Thr ^{ACC}
		Leu ^{CTT}	Leu ^{CTC}	Thr ^{ACT}	Thr ^{ACA}
Asn ^{AAT}	Asn ^{AAC}	Leu ^{CTA}	Leu ^{CTG}	Thr ^{ACA}	Thr ^{ACG}
		Leu ^{CTA}	Leu ^{CTC}	Thr ^{ACA}	Thr ^{ACC}
Asp ^{GAT}	Asp ^{GAC}			Thr ^{ACC}	Thr ^{ACG}
		Phe ^{TTC}	Phe ^{TTT}		
Cys ^{TGT}	Cys ^{TGC}			Tyr ^{TAT}	Tyr ^{TAC}
		Pro ^{CCG}	Pro ^{CCC}		
Glu ^{GAG}	Glu ^{GAA}	Pro ^{CCG}	Pro ^{CCT}	Val ^{GTA}	Val ^{GTG}
		Pro ^{CCA}	Pro ^{CCC}	Val ^{GTA}	Val ^{GTC}
Gly ^{GGC}	Gly ^{GGA}	Pro ^{CCA}	Pro ^{CCT}	Val ^{GTA}	Val ^{GTT}
Gly ^{GGT}	Gly ^{GGA}	Pro ^{CCT}	Pro ^{CCC}	Val ^{GTT}	Val ^{GTG}
Gly ^{GGG}	Gly ^{GGA}			Val ^{GTT}	Val ^{GTC}

Por el contrario, puede conseguirse una respuesta inmunitaria más débil o reducida al antígeno diana (por ejemplo, una respuesta inmunitaria que sea al menos de aproximadamente el 90 %, el 80 %, el 70 %, el 60 %, el 50 %, el 40 %, el 30 %, el 20 %, el 10 %, el 1 % y todos los porcentajes enteros entre medias, de la que se produce a partir del polinucleótido parental en condiciones idénticas) mediante la selección de un codón sinónimo que tenga una preferencia de respuesta inmunitaria menor que el primer codón que reemplaza. En realizaciones específicas de este tipo, el codón sinónimo se selecciona de manera que tenga una preferencia de respuesta inmunitaria que sea inferior a aproximadamente el 90 % de la preferencia de respuesta inmunitaria del codón que reemplaza. En ejemplos ilustrativos, los codones primero y sinónimo se seleccionan de la TABLA 5:

TABLA 5

Primer codón	Codón sinónimo	Primer codón	Codón sinónimo	Primer codón	Codón sinónimo
Ala ^{GCT}	Ala ^{GCG}	Ile ^{ATC}	Ile ^{ATA}	Ser ^{TCG}	Ser ^{TCT}
Ala ^{GCT}	Ala ^{GCA}	Ile ^{ATC}	Ile ^{ATT}	Ser ^{TCG}	Ser ^{TCA}
Ala ^{GCT}	Ala ^{GCC}	Ile ^{ATT}	Ile ^{ATA}	Ser ^{TCG}	Ser ^{TCC}
Ala ^{GCC}	Ala ^{GCG}			Ser ^{TCG}	Ser ^{AGC}
Ala ^{GCC}	Ala ^{GCA}	Leu ^{CTG}	Leu ^{CTA}	Ser ^{TCG}	Ser ^{AGT}
		Leu ^{CTG}	Leu ^{CTT}	Ser ^{TCT}	Ser ^{AGC}
Arg ^{CGA}	Arg ^{AGG}	Leu ^{CTG}	Leu ^{TTG}	Ser ^{TCT}	Ser ^{AGT}
Arg ^{CGA}	Arg ^{CGG}	Leu ^{CTG}	Leu ^{TTA}	Ser ^{TCA}	Ser ^{AGC}
Arg ^{CGC}	Arg ^{AGG}	Leu ^{CTC}	Leu ^{CTA}	Ser ^{TCA}	Ser ^{AGT}
Arg ^{CGC}	Arg ^{CGG}	Leu ^{CTC}	Leu ^{CTT}	Ser ^{TCC}	Ser ^{AGC}
Arg ^{CGT}	Arg ^{AGG}	Leu ^{CTC}	Leu ^{TTG}	Ser ^{TCC}	Ser ^{AGT}
Arg ^{CGT}	Arg ^{CGG}	Leu ^{CTC}	Leu ^{TTA}		
Arg ^{AGA}	Arg ^{AGG}	Leu ^{CTA}	Leu ^{TTG}	Thr ^{ACG}	Thr ^{ACC}
Arg ^{AGA}	Arg ^{CGG}	Leu ^{CTA}	Leu ^{TTA}	Thr ^{ACG}	Thr ^{ACA}
		Leu ^{CTT}	Leu ^{TTG}	Thr ^{ACG}	Thr ^{ACT}
Asn ^{AAC}	Asn ^{AAT}	Leu ^{CTT}	Leu ^{TTA}	Thr ^{ACC}	Thr ^{ACA}
		Leu ^{TTG}	Leu ^{TTA}	Thr ^{ACC}	Thr ^{ACT}
Asp ^{GAC}	Asp ^{GAT}			Thr ^{ACA}	Thr ^{ACT}
		Phe ^{TTT}	Phe ^{TTC}		
Cys ^{TGC}	Cys ^{TGT}			Tyr ^{TAC}	Tyr ^{TAT}
		Pro ^{CCC}	Pro ^{CCT}		
Glu ^{GAA}	Glu ^{GAG}	Pro ^{CCC}	Pro ^{CCA}	Val ^{GTG}	Val ^{GTT}
		Pro ^{CCC}	Pro ^{CCG}	Val ^{GTG}	Val ^{GTA}
Gly ^{GGA}	Gly ^{GGC}	Pro ^{CCT}	Pro ^{CCA}	Val ^{GTC}	Val ^{GTT}
Gly ^{GGA}	Gly ^{GGT}	Pro ^{CCT}	Pro ^{CCG}	Val ^{GTC}	Val ^{GTA}
Gly ^{GGA}	Gly ^{GGG}			Val ^{GTT}	Val ^{GTA}

En otros ejemplos ilustrativos, los codones primero y sinónimo se seleccionan de la TABLA 6:

TABLA 6

Primer codón	Codón sinónimo	Primer codón	Codón sinónimo	Primer codón	Codón sinónimo
Ala ^{GCT}	Ala ^{GCG}	Leu ^{CTA}	Leu ^{TTG}	Ser ^{TCG}	Ser ^{TCT}
Ala ^{GCT}	Ala ^{GCA}	Leu ^{CTA}	Leu ^{TTA}	Ser ^{TCG}	Ser ^{TCA}
Ala ^{GCT}	Ala ^{GCC}	Leu ^{CTT}	Leu ^{TTG}	Ser ^{TCG}	Ser ^{TCC}
		Leu ^{CTT}	Leu ^{TTA}	Ser ^{TCG}	Ser ^{AGC}
Arg ^{CGA}	Arg ^{AGG}	Leu ^{TTG}	Leu ^{TTA}	Ser ^{TCG}	Ser ^{AGT}
Arg ^{CGA}	Arg ^{CGG}			Ser ^{TCT}	Ser ^{AGC}
Arg ^{CGT}	Arg ^{AGG}	Phe ^{TTT}	Phe ^{TTC}	Ser ^{TCT}	Ser ^{AGT}
Arg ^{CGT}	Arg ^{CGG}			Ser ^{TCA}	Ser ^{AGC}
Arg ^{AGA}	Arg ^{AGG}	Pro ^{CCT}	Pro ^{CCA}	Ser ^{TCA}	Ser ^{AGT}
Arg ^{AGA}	Arg ^{CGG}	Pro ^{CCT}	Pro ^{CCG}	Ser ^{TCC}	Ser ^{AGC}
Glu ^{GAA}	Glu ^{GAG}			Thr ^{ACG}	Thr ^{ACC}
				Thr ^{ACG}	Thr ^{ACA}
Gly ^{GGA}	Gly ^{GGC}			Thr ^{ACG}	Thr ^{ACT}
Gly ^{GGA}	Gly ^{GGT}			Thr ^{ACA}	Thr ^{ACT}
Gly ^{GGA}	Gly ^{GGG}				
				Val ^{GTT}	Val ^{GTA}

Convenientemente, en algunos de los ejemplos ilustrativos indicados anteriormente, el método comprende adicionalmente seleccionar un segundo codón del polinucleótido parental para el reemplazo con un codón sinónimo, en el que el codón sinónimo se selecciona basándose en que presenta una preferencia de respuesta inmunitaria menor que el segundo codón en una comparación de preferencias de respuesta inmunitaria; y; (b) reemplazar el segundo codón con el codón sinónimo, en el que la comparación de preferencias de respuesta inmunitaria de los codones se representa mediante la TABLA 7:

TABLA 7

Segundo codón	Codón sinónimo	Segundo codón	Codón sinónimo	Segundo codón	Codón sinónimo
Ala ^{GCT}	Ala ^{GCG}	Ile ^{ATC}	Ile ^{ATA}	Ser ^{TCG}	Ser ^{TCT}
Ala ^{GCT}	Ala ^{GCA}	Ile ^{ATC}	Ile ^{ATT}	Ser ^{TCG}	Ser ^{TCA}
Ala ^{GCT}	Ala ^{GCC}	Ile ^{ATT}	Ile ^{ATA}	Ser ^{TCG}	Ser ^{TCC}
Ala ^{GCC}	Ala ^{GCG}			Ser ^{TCG}	Ser ^{AGC}
Ala ^{GCC}	Ala ^{GCA}	Leu ^{CTG}	Leu ^{CTA}	Ser ^{TCG}	Ser ^{AGT}
		Leu ^{CTG}	Leu ^{CTT}	Ser ^{TCT}	Ser ^{AGC}
Arg ^{CGA}	Arg ^{AGG}	Leu ^{CTG}	Leu ^{TTG}	Ser ^{TCT}	Ser ^{AGT}
Arg ^{CGA}	Arg ^{CGG}	Leu ^{CTG}	Leu ^{TTA}	Ser ^{TCA}	Ser ^{AGC}
Arg ^{CGC}	Arg ^{AGG}	Leu ^{CTC}	Leu ^{CTA}	Ser ^{TCA}	Ser ^{AGT}
Arg ^{CGC}	Arg ^{CGG}	Leu ^{CTC}	Leu ^{CTT}	Ser ^{TCC}	Ser ^{AGC}
Arg ^{CGT}	Arg ^{AGG}	Leu ^{CTC}	Leu ^{TTG}	Ser ^{TCC}	Ser ^{AGT}
Arg ^{CGT}	Arg ^{CGG}	Leu ^{CTC}	Leu ^{TTA}		
Arg ^{AGA}	Arg ^{AGG}	Leu ^{CTA}	Leu ^{TTG}	Thr ^{ACG}	Thr ^{ACC}
Arg ^{AGA}	Arg ^{CGG}	Leu ^{CTA}	Leu ^{TTA}	Thr ^{ACG}	Thr ^{ACA}
		Leu ^{CTT}	Leu ^{TTG}	Thr ^{ACG}	Thr ^{ACT}
Asn ^{AAC}	Asn ^{AAT}	Leu ^{CTT}	Leu ^{TTA}	Thr ^{ACC}	Thr ^{ACA}
		Leu ^{TTG}	Leu ^{TTA}	Thr ^{ACC}	Thr ^{ACT}
Asp ^{GAC}	Asp ^{GAT}			Thr ^{ACA}	Thr ^{ACT}
		Phe ^{TTT}	Phe ^{TTC}		
Cys ^{TGC}	Cys ^{TGT}			Tyr ^{TAC}	Tyr ^{TAT}
		Pro ^{CCC}	Pro ^{CCT}		
Glu ^{GAA}	Glu ^{GAG}	Pro ^{CCC}	Pro ^{CCA}	Val ^{GTG}	Val ^{GTT}
		Pro ^{CCC}	Pro ^{CCG}	Val ^{GTG}	Val ^{GTA}
Gly ^{GGA}	Gly ^{GGC}	Pro ^{CCT}	Pro ^{CCA}	Val ^{GTC}	Val ^{GTT}
Gly ^{GGA}	Gly ^{GGT}	Pro ^{CCT}	Pro ^{CCG}	Val ^{GTC}	Val ^{GTA}
Gly ^{GGA}	Gly ^{GGG}			Val ^{GTT}	Val ^{GTA}

- 10 En otro aspecto más, la divulgación proporciona un polinucleótido sintético construido de acuerdo con uno cualquiera de los métodos anteriores.

- De acuerdo con la presente invención, los polinucleótidos sintéticos que se construyen mediante los métodos que se describen en el presente documento son útiles para la expresión en un mamífero para desencadenar una respuesta inmunitaria a un antígeno diana. En consecuencia, en otro aspecto más, la presente divulgación proporciona construcciones químicas que comprenden un polinucleótido sintético de la invención, que está conectado operativamente a una secuencia reguladora.

- En algunas realizaciones, la construcción química está en forma de una composición farmacéutica que comprende opcionalmente un excipiente y/o vehículo farmacéuticamente aceptable. En consecuencia, en otro aspecto, la divulgación proporciona composiciones farmacéuticas que son útiles para modular una respuesta inmunitaria a un antígeno diana en un mamífero, respuesta conferida por la expresión de un polinucleótido parental que codifica un polipéptido que corresponde a al menos una porción del antígeno diana. Estas composiciones en general comprenden una construcción química y un excipiente y/o vehículo farmacéuticamente aceptable, en las que la construcción química comprende un polinucleótido sintético que está conectado operativamente a una secuencia reguladora y que se distingue del polinucleótido parental por el reemplazo de un primer codón en el polinucleótido parental por un codón sinónimo que tiene una preferencia de respuesta inmunitaria diferente a la del primer codón y en las que los codones primero y sinónimo se seleccionan de acuerdo con una cualquiera de las TABLAS 2, 3, 5 y 6. En algunas realizaciones, las composiciones comprenden adicionalmente un adyuvante que potencia la eficacia de la respuesta inmunitaria. En algunas realizaciones, la composición se formula para la administración transcutánea o

dérmica, por ejemplo, mediante entrega biolística o por microaguja o por inyección intradérmica. Convenientemente, en realizaciones en las que se desea una respuesta inmunitaria más fuerte o potenciada al antígeno diana, los codones primero y sinónimo se seleccionan de acuerdo con las TABLAS 2 o 3. Por el contrario, en realizaciones en las que se desea una respuesta inmunitaria más débil o reducida al antígeno diana, los codones primero y sinónimo se seleccionan de acuerdo con las TABLAS 5 o 6.

En otro aspecto más, la divulgación abarca métodos de modulación de la cualidad de una respuesta inmunitaria a un antígeno diana en un mamífero, respuesta conferida por la expresión de un polinucleótido parental que codifica un polipéptido que corresponde a al menos una porción del antígeno diana. Estos métodos en general comprenden: introducir en el mamífero un polinucleótido sintético que está conectado operativamente a una secuencia reguladora y que se distingue del polinucleótido parental por el reemplazo de un primer codón en el polinucleótido parental por un codón sinónimo que tiene una preferencia de respuesta inmunitaria diferente de la del primer codón y en los que los codones primero y sinónimo se seleccionan de acuerdo con una cualquiera de las TABLAS 2, 3, 5 y 6. En estos métodos, la expresión del polinucleótido sintético da como resultado una cualidad de respuesta inmunitaria diferente (por ejemplo, más fuerte o más débil) de la obtenida a través de la expresión del polinucleótido parental en las mismas condiciones. Convenientemente, la construcción quimérica se introduce en el mamífero mediante la entrega de la construcción a las células presentadoras de antígeno (por ejemplo, células dendríticas, macrófagos, células de Langerhans o sus precursores) del mamífero. En algunas realizaciones, la construcción quimérica se introduce en la dermis y/o epidermis del mamífero (por ejemplo, administración transcutánea o intradérmica) y, en este sentido, se prevé cualquier sitio de administración adecuado, incluyendo el abdomen. Generalmente, la respuesta inmunitaria se selecciona entre una respuesta mediada por células y una respuesta inmunitaria humoral. En realizaciones específicas, la respuesta inmunitaria es una respuesta inmunitaria humoral.

En un aspecto relacionado, la divulgación abarca métodos de potenciación de la cualidad de una respuesta inmunitaria a un antígeno diana en un mamífero, respuesta conferida por la expresión de un polinucleótido parental que codifica un polipéptido que corresponde a al menos una porción del antígeno diana. Estos métodos en general comprenden: introducir en el mamífero una construcción quimérica que comprende un polinucleótido sintético que está conectado operativamente a una secuencia reguladora y que se distingue del polinucleótido parental por el reemplazo de un primer codón en el polinucleótido parental por un codón sinónimo que tiene una preferencia de respuesta inmunitaria mayor que el primer codón, en los que los codones primero y sinónimo se seleccionan de acuerdo con las TABLAS 2 o 3. En estos métodos, la expresión del polinucleótido sintético normalmente da como resultado una respuesta inmunitaria más fuerte o potenciada que la que se obtiene a través de la expresión del polinucleótido parental en las mismas condiciones.

En otro aspecto relacionado, la divulgación se extiende a métodos de disminución de la cualidad de una respuesta inmunitaria a un antígeno diana en un mamífero, respuesta conferida por la expresión de un polinucleótido parental que codifica un polipéptido que corresponde a al menos una porción del antígeno diana. Estos métodos en general comprenden: introducir en el mamífero una construcción quimérica que comprende un polinucleótido sintético que está conectado operativamente a una secuencia reguladora y que se distingue del polinucleótido parental por el reemplazo de un primer codón en el polinucleótido parental por un codón sinónimo que tiene una preferencia de respuesta inmunitaria menor que el primer codón, en los que los codones primero y sinónimo se seleccionan de acuerdo con las TABLAS 5 o 6. En estos métodos, la expresión del polinucleótido sintético normalmente da como resultado una respuesta inmunitaria más débil o reducida que la que se obtiene a través de la expresión de la polinucleótido parental en las mismas condiciones.

Sin embargo, un aspecto adicional de la presente divulgación abarca métodos de potenciación de la cualidad de una respuesta inmunitaria a un antígeno diana en un mamífero, respuesta conferida por la expresión de un primer polinucleótido que codifica un polipéptido que corresponde a al menos una porción del antígeno diana. Estos métodos en general comprenden: reintroducir en el mamífero una primera construcción de ácido nucleico que comprende el primer polinucleótido en conexión operable con una secuencia reguladora; y una segunda construcción de ácido nucleico que comprende un segundo polinucleótido que está conectado operativamente a una secuencia reguladora y que codifica un iso-ARNt que corresponde a un codón del primer polinucleótido, en el que el codón tiene una preferencia de respuesta inmunitaria baja o intermedia y se selecciona entre el grupo que consiste en Ala^{GCA}, Ala^{GCG}, Ala^{GCC}, Arg^{AGG}, Arg^{CGG}, Asn^{AAT}, Asp^{GAT}, Cys^{TGT}, Glu^{GAG}, Gly^{GGG}, Gly^{GGT}, Gly^{GGC}, Ile^{ATA}, Ile^{ATT}, Leu^{TTG}, Leu^{TTA}, Leu^{CTA}, Leu^{CTT}, Phe^{TTC}, Pro^{CCA}, Pro^{CCG}, Pro^{CCT}, Ser^{AGC}, Ser^{AGT}, Ser^{TCT}, Ser^{TCA}, Ser^{TCC}, Thr^{ACA}, Thr^{ACT}, Tyr^{TAT}, Val^{GTA} y Val^{GTT}. En realizaciones específicas, el codón tiene una preferencia de respuesta inmunitaria 'baja' y se selecciona entre el grupo que consiste en Ala^{GCA}, Ala^{GCG}, Arg^{AGG}, Arg^{CGG}, Asn^{AAT}, Asp^{GAT}, Cys^{TGT}, Glu^{GAG}, Gly^{GGG}, Gly^{GGT}, Gly^{GGC}, Ile^{ATA}, Leu^{TTG}, Leu^{TTA}, Phe^{TTC}, Pro^{CCA}, Pro^{CCG}, Ser^{AGC}, Ser^{AGT}, Thr^{TAT}, Tyr^{TAT} y Val^{GTA}.

Breve descripción de los dibujos

La Figura 1 es una representación esquemática que representa una alineación de secuencia de nucleótidos de construcciones de E7 de ALA secretadas y controles (IgkC1, IgkS1-1, IgkS1-2, IgkS1-3, IgkS1-4 y IgkC2) como se definen adicionalmente en el Ejemplo 1 y en la Tabla 12. Las secuencias se unen en los sitios *KpnI* y *EcoRI* de pCDNA3.

La Figura 2 es una representación esquemática que representa una alineación de secuencia de nucleótidos de construcciones de E7 de ARG secretadas y controles (IgkS1-5, IgkS1-6, IgkS1-7, IgkS1-8, IgkS1-9, IgkS1-10, IgkC1 y IgkC2) como se definen adicionalmente en el Ejemplo 1 y en la Tabla 12. Las secuencias se unen en los sitios *KpnI* y *EcoRI* de pCDNA3.

La Figura 3 es una representación esquemática que representa una alineación de secuencia de nucleótidos de construcciones de E7 de ASN y LYS secretadas y controles (IgkC1, IgkS1-12, IgkS1-31 y IgkC2) como se define adicionalmente en el Ejemplo 1 y en la Tabla 12. Las secuencias se unen en los sitios *KpnI* y *EcoRI* de pCDNA3.

La Figura 4 es una representación esquemática que representa una alineación de secuencia de nucleótidos de construcciones de E7 de ASP secretadas y controles (IgkC1, IgkS1-13, IgkS1-14 y IgkC2) como se define adicionalmente en el Ejemplo 1 y en la Tabla 12. Las secuencias se unen en los sitios *KpnI* y *EcoRI* de pCDNA3.

La Figura 5 es una representación esquemática que representa una alineación de secuencia de nucleótidos de construcciones de E7 de CYS secretadas y controles (IgkC1, IgkS1-15, IgkS1-16 y IgkC2) como se define adicionalmente en el Ejemplo 1 y en la Tabla 12. Las secuencias se unen en los sitios *KpnI* y *EcoRI* de pCDNA3.

La Figura 6 es una representación esquemática que representa una alineación de secuencia de nucleótidos de construcciones de E7 de GLU secretadas y controles (IgkS1-17, IgkS1-18, IgkC2 y IgkC1) como se define adicionalmente en el Ejemplo 1 y en la Tabla 12. Las secuencias se unen en los sitios *KpnI* y *EcoRI* de pCDNA3.

La Figura 7 es una representación esquemática que representa una alineación de secuencia de nucleótidos de construcciones de E7 de GLN secretadas y controles (IgkC1, IgkS1-19, IgkS1-20 y IgkC2) como se define adicionalmente en el Ejemplo 1 y en la Tabla 12. Las secuencias se unen en los sitios *KpnI* y *EcoRI* de pCDNA3.

La Figura 8 es una representación esquemática que representa una alineación de secuencia de nucleótidos de construcciones de E7 de GLY secretadas y controles (IgkC1, IgkS1-21, IgkS1-22, IgkS1-23, IgkS1-24 y IgkC2) como se define adicionalmente en el Ejemplo 1 y la Tabla 12. Las secuencias se unen en los sitios *KpnI* y *EcoRI* de pCDNA3.

La Figura 9 es una representación esquemática que representa una alineación de secuencia de nucleótidos de construcciones de E7 de HIS secretadas y controles (IgkC1, IgkS1-25, IgkS1-26 y IgkC2) como se define adicionalmente en el Ejemplo 1 y en la Tabla 12. Las secuencias se unen en los sitios *KpnI* y *EcoRI* de pCDNA3.

La Figura 10 es una representación esquemática que representa una alineación de secuencia de nucleótidos de construcciones de E7 de ILE secretadas y controles (IgkC1, IgkS1-27, IgkS1-28, IgkS1-29 y IgkC2) como se define adicionalmente en el Ejemplo 1 y en la Tabla 12. Las secuencias se unen en los sitios *KpnI* y *EcoRI* de pCDNA3.

La Figura 11 es una representación esquemática que representa una alineación de secuencia de nucleótidos de construcciones de E7 de LEU secretadas y controles (IgkS1-50, IgkS1-51, IgkS1-52, IgkS1-53, IgkS1-54, IgkS1-55, IgkC3 y IgkC4) como se define adicionalmente en el Ejemplo 1 y en la Tabla 12. Las secuencias se unen en los sitios *KpnI* y *EcoRI* de pCDNA3. Las construcciones de E7 de LEU son oncogénicas (es decir, codifican proteína E7 de tipo silvestre).

La Figura 12 es una representación esquemática que representa una alineación de secuencia de nucleótidos de construcciones de E7 de PHE secretadas y controles (IgkS1-32, IgkS1-33, IgkC1 y IgkC2) como se define adicionalmente en el Ejemplo 1 y en la Tabla 12. Las secuencias se unen en los sitios *KpnI* y *EcoRI* de pCDNA3. Los dos restos de LEU se mutaron a PHE en esta secuencia de manera que hubiera tres en lugar de un resto de PHE.

La Figura 13 es una representación esquemática que representa una alineación de secuencia de nucleótidos de construcciones de E7 de PRO secretadas y controles (IgkS1-56, IgkS1-57, IgkS1-58, IgkS1-59, IgkC3 y IgkC4) como se definen adicionalmente en el Ejemplo 1 y la Tabla 12. Las secuencias se unen en los sitios *KpnI* y *EcoRI* de pCDNA3. Las construcciones de E7 de PRO son oncogénicas (es decir, codifican proteína E7 de tipo silvestre).

La Figura 14 es una representación esquemática que representa una alineación de secuencia de nucleótidos de construcciones de E7 de SER secretadas y controles (IgkS1-34, IgkS1-35, IgkS1-36, IgkS1-37, IgkS1-38, IgkS1-39, IgkC1 y IgkC2) como se definen adicionalmente en el Ejemplo 1 y en la Tabla 12. Las secuencias se unen en los sitios *KpnI* y *EcoRI* de pCDNA3.

La Figura 15 es una representación esquemática que representa una alineación de secuencia de nucleótidos de construcciones de E7 de THR secretadas y controles (IgkC1, IgkS1-40, IgkS1-41, IgkS1-42, IgkS1-43 y IgkC2) como se definen adicionalmente en el Ejemplo 1 y la Tabla 12. Las secuencias se unen en los sitios *KpnI* y *EcoRI* de pCDNA3.

La Figura 16 es una representación esquemática que representa una alineación de secuencia de nucleótidos de construcciones de E7 de TYR secretadas y controles (IgkC1, IgkS1-44, IgkS1-45 y IgkC2) como se definen adicionalmente en el Ejemplo 1 y en la Tabla 12. Las secuencias se unen en los sitios *KpnI* y *EcoRI* de pCDNA3.

La Figura 17 es una representación esquemática que representa una alineación de secuencia de nucleótidos de construcciones de E7 de VAL secretadas y controles (IgkC1, IgkS1-46, IgkS1-47, IgkS1-48, IgkS1-49 y IgkC2) como se definen adicionalmente en el Ejemplo 1 y en la Tabla 12. Las secuencias se unen en los sitios *KpnI* y *EcoRI* de pCDNA3.

La Figura 18 es una representación gráfica que muestra la respuesta a la inmunización por pistola de genes con construcciones de E7 optimizadas y desoptimizadas medida mediante (a) ELISA, (b) ELISPOT de células de B de

memoria y (c) ELISPOT de IFN- γ . Para la parte (a) se inmunizaron ocho ratones por grupo (4 inmunizaciones, 3 semanas de diferencia) y el suero se tomó tres semanas después de la inmunización final; (izquierda) ELISA de proteína E7, (derecha) ELISA 101 de péptido E7. Los pocillos se realizaron por duplicado. Para las partes (b) y (c) los ratones se inmunizaron dos veces, con tres semanas de diferencia y los bazos se recolectaron tres semanas después de la segunda inmunización. Los bazos se agruparon antes del análisis. Las ELISPOT de células de memoria B e IFN- γ se realizaron dos veces y tres veces, respectivamente, y los pocillos se realizaron por triplicado. Se usaron tres ratones por grupo por repetición. Los resultados mostrados en las partes (b) y (c) son de experimentos individuales y son representativos de los conjuntos de datos completos. Los datos experimentales de ELISPOT particulares incluidos en el presente documento se juntaron con los datos correspondientes en la Figura 20 y, por tanto, pueden compararse directamente. Se usaron ensayos t de dos colas no emparejadas para comparar las construcciones modificadas con las de tipo silvestre. *** $P < 0,001$, ** $0,001 \leq P < 0,01$, * $0,01 \leq P \leq 0,05$, ns = no significativo ($P > 0,05$). En (a) O1-O3 no fueron significativamente diferentes de MC según se midió mediante ensayos t de dos colas no emparejados. ts = E7 de uso de codones de tipo silvestre; O1-O3 = construcciones de E7 optimizadas por codones 1 a 3; W = codón desoptimizado E7; CM = E7 de uso de codones de consenso de mamíferos.

La Figura 19 es una representación gráfica que representa la respuesta a la inmunización mediante inyección intradérmica con construcciones optimizadas y desoptimizadas medida mediante (a) ELISA, (b) ELISPOT de células de B de memoria y (c) ELISPOT de IFN- γ . Para la parte (a) se inmunizaron ocho ratones por grupo (4 inmunizaciones, 3 semanas de diferencia) y el suero se tomó tres semanas después de la inmunización final; (izquierda) ELISA de proteína E7, (derecha) ELISA 101 de péptido E7. Los pocillos se realizaron por duplicado. Para las partes (b) y (c) los ratones se inmunizaron dos veces, con tres semanas de diferencia y los bazos se recolectaron tres semanas después de la segunda inmunización. Los bazos se agruparon antes del análisis. Las ELISPOT de células de memoria B e IFN- γ se realizaron dos veces y tres veces, respectivamente, y los pocillos se realizaron por triplicado. Se usaron tres ratones por grupo por repetición. Los resultados mostrados en las partes (b) y (c) son de experimentos individuales y son representativos de los conjuntos de datos completos. Los datos experimentales de ELISPOT particulares incluidos en el presente documento se juntaron con los datos correspondientes en la Figura 20 y, por tanto, pueden compararse directamente. Se usaron ensayos t de dos colas no emparejadas para comparar las construcciones modificadas con las de tipo silvestre. *** $P < 0,001$, ** $0,001 \leq P < 0,01$, * $0,01 \leq P \leq 0,05$, ns = no significativo ($P > 0,05$). En (a) O1-O3 no fueron significativamente diferentes de MC según se midió mediante ensayos t de dos colas no emparejados. ts = E7 de uso de codones de tipo silvestre; O1-O3 = construcciones de E7 optimizadas por codones 1 a 3; W = codón desoptimizado E7; CM = E7 de uso de codones de consenso de mamíferos.

La Figura 20 es una representación gráfica que muestra los resultados de un ELISA que mide la unión de suero de ratones inmunizados con diversas construcciones gD2 mediante inyección intradérmica (barras de color blanco) o inmunización por pistola de genes (barras de color negro), a gD2tr marcada con His C-terminal. Nótese que la proteína gD2tr marcada con His se usó en un estado no purificado (en el sobrenadante de células CHO) y que las lecturas de fondo de unión no específica a sobrenadante de control se han restado de los resultados.

TABLA 8

BREVE DESCRIPCIÓN DE LAS SECUENCIAS		
NÚMERO DE ID DE SECUENCIA	SECUENCIA	LONGITUD
SEQ ID NO: 1	Secuencia de nucleótidos de la construcción IgkS2-13 Asp GAT	387 nucleótidos
SEQ ID NO: 2	Secuencia de nucleótidos de la construcción IgkS2-14 Asp GAC	387 nucleótidos
SEQ ID NO: 3	Secuencia de nucleótidos de la construcción IgkS2-15 Cys TGT	387 nucleótidos
SEQ ID NO: 4	Secuencia de nucleótidos de la construcción IgkS2-16 Cys TGC	387 nucleótidos
SEQ ID NO: 5	Secuencia de nucleótidos de la construcción IgkS2-17 Glu GAG	387 nucleótidos
SEQ ID NO: 6	Secuencia de nucleótidos de la construcción IgkS2-18 Glu GAA	387 nucleótidos
SEQ ID NO: 7	Secuencia de nucleótidos de la construcción IgkS2-19 Gln CAG	387 nucleótidos
SEQ ID NO: 8	Secuencia de nucleótidos de la construcción IgkS2-20 Gln CAA	387 nucleótidos
SEQ ID NO: 9	Secuencia de nucleótidos de la construcción IgkS2-21 Gly GGG	387 nucleótidos
SEQ ID NO: 10	Secuencia de nucleótidos de la construcción IgkS2-22 Gly GGA	387 nucleótidos
SEQ ID NO: 11	Secuencia de nucleótidos de la construcción IgkS2-23 Gly GGT	387 nucleótidos
SEQ ID NO: 12	Secuencia de nucleótidos de la construcción IgkS2-24 Gly GGC	387 nucleótidos
SEQ ID NO: 13	Secuencia de nucleótidos de la construcción IgkS2-27 Ile ATA	387 nucleótidos
SEQ ID NO: 14	Secuencia de nucleótidos de la construcción IgkS2-28 Ile ATT	387 nucleótidos
SEQ ID NO: 15	Secuencia de nucleótidos de la construcción IgkS2-29 Ile ATC	387 nucleótidos
SEQ ID NO: 16	Secuencia de nucleótidos de la construcción IgkS2-34 Ser AGT	387 nucleótidos

BREVE DESCRIPCIÓN DE LAS SECUENCIAS		
NÚMERO DE ID DE SECUENCIA	SECUENCIA	LONGITUD
SEQ ID NO: 17	Secuencia de nucleótidos de la construcción IgkS2-35 Ser AGC	387 nucleótidos
SEQ ID NO: 18	Secuencia de nucleótidos de la construcción IgkS2-36 Ser TCG	387 nucleótidos
SEQ ID NO: 19	Secuencia de nucleótidos de la construcción IgkS2-37 Ser TCA	387 nucleótidos
SEQ ID NO: 20	Secuencia de nucleótidos de la construcción IgkS2-38 Ser TCT	387 nucleótidos
SEQ ID NO: 21	Secuencia de nucleótidos de la construcción IgkS2-39 Ser TCC	387 nucleótidos
SEQ ID NO: 22	Secuencia de nucleótidos de la construcción IgkS2-40 Thr ACG	387 nucleótidos
SEQ ID NO: 23	Secuencia de nucleótidos de la construcción IgkS2-41 Thr ACA	387 nucleótidos
SEQ ID NO: 24	Secuencia de nucleótidos de la construcción IgkS2-42 Thr ACT	387 nucleótidos
SEQ ID NO: 25	Secuencia de nucleótidos de la construcción IgkS2-43 Thr ACC	387 nucleótidos
SEQ ID NO: 26	Secuencia de nucleótidos de la construcción IgkS2-46 Val GTG	387 nucleótidos
SEQ ID NO: 27	Secuencia de nucleótidos de la construcción IgkS2-47 Val GTA	387 nucleótidos
SEQ ID NO: 28	Secuencia de nucleótidos de la construcción IgkS2-48 Val GTT	387 nucleótidos
SEQ ID NO: 29	Secuencia de nucleótidos de la construcción IgkS2-49 Val GTG	387 nucleótidos
SEQ ID NO: 30	Secuencia de nucleótidos del enlazador IgkS2-1 Ala GCG	408 nucleótidos
SEQ ID NO: 31	Secuencia de nucleótidos del enlazador IgkS2-2 Ala GCA	408 nucleótidos
SEQ ID NO: 32	Secuencia de nucleótidos del enlazador IgkS2-3 Ala GCT	408 nucleótidos
SEQ ID NO: 33	Secuencia de nucleótidos del enlazador IgkS2-4 Ala GCC	408 nucleótidos
SEQ ID NO: 34	Secuencia de nucleótidos del enlazador IgkS2-5 Arg AGG	408 nucleótidos
SEQ ID NO: 35	Secuencia de nucleótidos del enlazador IgkS2-6 Arg AGA	408 nucleótidos
SEQ ID NO: 36	Secuencia de nucleótidos del enlazador IgkS2-7 Arg CGG	408 nucleótidos
SEQ ID NO: 37	Secuencia de nucleótidos del enlazador IgkS2-8 Arg CGA	408 nucleótidos
SEQ ID NO: 38	Secuencia de nucleótidos del enlazador IgkS2-9 Arg CGT	408 nucleótidos
SEQ ID NO: 39	Secuencia de nucleótidos del enlazador IgkS2-10 Arg CGC	408 nucleótidos
SEQ ID NO: 40	Secuencia de nucleótidos del enlazador IgkS2-11 Asn AAT	408 nucleótidos
SEQ ID NO: 41	Secuencia de nucleótidos del enlazador IgkS2-12 Asn AAC	408 nucleótidos
SEQ ID NO: 42	Secuencia de nucleótidos del enlazador IgkS2-25 His CAT	408 nucleótidos
SEQ ID NO: 43	Secuencia de nucleótidos del enlazador IgkS2-26 His CAC	408 nucleótidos
SEQ ID NO: 44	Secuencia de nucleótidos del enlazador IgkS2-30 Lys AAG	408 nucleótidos
SEQ ID NO: 45	Secuencia de nucleótidos del enlazador IgkS2-31 Lys AAA	408 nucleótidos
SEQ ID NO: 46	Secuencia de nucleótidos del enlazador IgkS2-32 Phe TTT	408 nucleótidos
SEQ ID NO: 47	Secuencia de nucleótidos del enlazador IgkS2-33 Phe TTC	408 nucleótidos
SEQ ID NO: 48	Secuencia de nucleótidos del enlazador IgkS2-44 Tyr TAT	408 nucleótidos
SEQ ID NO: 49	Secuencia de nucleótidos del enlazador IgkS2-45 Tyr TAC	408 nucleótidos
SEQ ID NO: 50	Hemaglutinina HA del Virus de la Gripe A (A/Hong Kong/213/03(H5N1)) BAE07201 de tipo silvestre	1707 nucleótidos
SEQ ID NO: 51	Hemaglutinina HA del Virus de la Gripe A (A/Hong Kong/213/03(H5N1)) BAE07201 de tipo silvestre	568 aa
SEQ ID NO: 52	Hemaglutinina HA del Virus de la Gripe A (A/Hong Kong/213/03(H5N1)) Modificada por codones	1707 nucleótidos
SEQ ID NO: 53	Hemaglutinina HA del Virus de la Gripe A (A/cerdo/Corea/PZ72-1/2006 (H3N1)) DQ923506 de tipo silvestre	1701 nucleótidos
SEQ ID NO: 54	Hemaglutinina HA del Virus de la Gripe A (A/cerdo/Corea/PZ72-1/2006 (H3N1)) DQ923506 de tipo silvestre	566 aa
SEQ ID NO: 55	Hemaglutinina HA del Virus de la Gripe A (A/cerdo/Corea/PZ72-1/2006 (H3N1)) Modificada por codones	1701 nucleótidos

BREVE DESCRIPCIÓN DE LAS SECUENCIAS		
NÚMERO DE ID DE SECUENCIA	SECUENCIA	LONGITUD
SEQ ID NO: 56	Neuraminidasa NA del Virus de la Gripe A (A/Hong Kong/213/03(H5N1)) AB212056 de tipo silvestre	1410 nucleótidos
SEQ ID NO: 57	Neuraminidasa NA del Virus de la Gripe A (A/Hong Kong/213/03(H5N1)) AB212056 de tipo silvestre	469 aa
SEQ ID NO: 58	Neuraminidasa NA del Virus de la Gripe A (A/Hong Kong/213/03(H5N1)) Modificada por codones	1410 nucleótidos
SEQ ID NO: 59	Neuraminidasa NA del Virus de la Gripe A (A/cerdo/MI/PU243/04 (H3N1)) DQ150427 de tipo silvestre	1410 nucleótidos
SEQ ID NO: 60	Neuraminidasa NA del Virus de la Gripe A (A/cerdo/MI/PU243/04 (H3N1)) DQ150427 de tipo silvestre	469 aa
SEQ ID NO: 61	Neuraminidasa NA del Virus de la Gripe A (A/cerdo/MI/PU243/04 (H3N1)) Modificada por codones	1410 nucleótidos
SEQ ID NO: 62	E1 del Virus de la Hepatitis C (Serotipo 1A, aislado H77) AF009606 de tipo silvestre	576 nucleótidos
SEQ ID NO: 63	E1 del Virus de la Hepatitis C (Serotipo 1A, aislado H77) NP 751920 de tipo silvestre	192 aa
SEQ ID NO: 64	E1 del Virus de la Hepatitis C (Serotipo 1A, aislado H77) Modificado por codones	576 nucleótidos
SEQ ID NO: 65	E2 del Virus de la Hepatitis C (Serotipo 1A, aislado H77) AF009606 de tipo silvestre	1089 nucleótidos
SEQ ID NO: 66	E2 del Virus de la Hepatitis C (Serotipo 1A, aislado H77) NP 751921 de tipo silvestre	363 aa
SEQ ID NO: 67	E2 del Virus de la Hepatitis C (Serotipo 1A, aislado H77) Modificado por codones	1089 nucleótidos
SEQ ID NO: 68	Virus de Epstein Barr (Tipo 1, gp350 B95-8) NC 007605 de tipo silvestre	2724 nucleótidos
SEQ ID NO: 69	Virus de Epstein Barr (Tipo 1, gp350 B95-8) CAD53417 de tipo silvestre	907 aa
SEQ ID NO: 70	Virus de Epstein Barr (Tipo 1, gp350 B95-8) Modificado por codones	2724 nucleótidos
SEQ ID NO: 71	Virus de Epstein Barr (Tipo 2, gp350 AG876) NC 009334 de tipo silvestre	2661 nucleótidos
SEQ ID NO: 72	Virus de Epstein Barr (Tipo 2, gp350 AG876) YP 001129462 de tipo silvestre	886 aa
SEQ ID NO: 73	Virus de Epstein Barr (Tipo 2, gp350 AG876) Modificado por codones	2661 nucleótidos
SEQ ID NO: 74	Virus del Herpes Simple 2 (Glucoproteína B cepa HG52) NC 001798 de tipo silvestre	2715 nucleótidos
SEQ ID NO: 75	Virus del Herpes Simple 2 (Glucoproteína B cepa HG52) CAB06752 de tipo silvestre	904 aa
SEQ ID NO: 76	Virus del Herpes Simple 2 (Glucoproteína B cepa HG52) Modificado por codones	2715 nucleótidos
SEQ ID NO: 77	Virus del Herpes Simple (Glucoproteína D cepa HG52) NC 001798 de tipo silvestre	1182 nucleótidos
SEQ ID NO: 78	Virus del Herpes Simple (Glucoproteína D cepa HG52) NP 0044536 de tipo silvestre	393 aa
SEQ ID NO: 79	Virus del Herpes Simple (Glucoproteína D cepa HG52) Modificado por codones	1182 nucleótidos
SEQ ID NO: 80	VPH-16 E7 de tipo silvestre	387 nucleótidos
SEQ ID NO: 81	VPH-16 E7 O1	387 nucleótidos
SEQ ID NO: 82	VPH-16 E7 O2	387 nucleótidos

BREVE DESCRIPCIÓN DE LAS SECUENCIAS		
NÚMERO DE ID DE SECUENCIA	SECUENCIA	LONGITUD
SEQ ID NO: 83	VPH-16 E7 O3	417 nucleótidos
SEQ ID NO: 84	VPH-16 E7 W	387 nucleótidos
SEQ ID NO: 85	VHS-2 gD2 de tipo silvestre	1182 nucleótidos
SEQ ID NO: 86	VHS-2 gD2 O1	1182 nucleótidos
SEQ ID NO: 87	VHS-2 gD2 O2	1182 nucleótidos
SEQ ID NO: 88	VHS-2 gD2 O3	1182 nucleótidos
SEQ ID NO: 89	VHS-2 gD2 W	1182 nucleótidos
SEQ ID NO: 90	Cebador directo común	41 nucleótidos
SEQ ID NO: 91	ODN-7909	24 nucleótidos

Descripción detallada de la invención

1. Definiciones

A menos que se defina lo contrario, todos los términos técnicos y científicos utilizados en el presente documento tienen el mismo significado que se entiende habitualmente por los expertos habituales en la materia a la que pertenece la invención. Aunque pueden usarse cualesquiera métodos y materiales similares o equivalentes a los descritos en el presente documento en la práctica o ensayo de la presente invención, se describen métodos y materiales preferidos. Para los fines de la presente invención, los siguientes términos se definen a continuación.

Los artículos "un" y "una" se usan en el presente documento para referirse a uno o a más de uno (es decir, a al menos uno) de los objetos gramaticales del artículo. A modo de ejemplo, "un elemento" significa un elemento o más de un elemento.

Por "aproximadamente" se quiere decir una cantidad, nivel, valor, frecuencia, porcentaje, dimensión, tamaño o cantidad que varía en no más del 15 % y, preferentemente, en no más del 10 %, el 9 %, el 8 %, el 7 %, el 6 %, el 5 %, el 4 %, el 3 %, el 2 %, el 1 % con respecto a una cantidad, nivel, valor, frecuencia, porcentaje, dimensión, tamaño o cantidad de referencia.

Las expresiones "administración simultánea" o "administrar simultáneamente" o "coadministración" y similares se refieren a la administración de una única composición que contiene dos o más principios activos o a la administración de cada principio activo como composiciones separadas y/o entregadas por vías separadas ya sea contemporáneamente o simultáneamente o secuencialmente en un período de tiempo lo suficientemente corto para que el resultado eficaz sea equivalente al obtenido cuando todos dichos principios activos se administran como una sola composición. Por "simultáneamente" se quiere decir que los agentes activos se administran sustancialmente al mismo tiempo y de manera deseable juntos en la misma formulación. Por "contemporáneamente" se quiere decir que los agentes activos se administran cercanos en el tiempo, por ejemplo, un agente se administra en el plazo de aproximadamente un minuto hasta el plazo de aproximadamente un día antes o después de otro. Cualquier momento contemporáneo es útil. Sin embargo, con frecuencia se dará el caso de que cuando no se administran de forma simultánea, los agentes se administrarán en el plazo de aproximadamente un minuto hasta el plazo de aproximadamente ocho horas y preferentemente en menos de aproximadamente uno a aproximadamente cuatro horas. Cuando se administran simultáneamente, los agentes se administran convenientemente en el mismo sitio en el sujeto. La expresión "el mismo sitio" incluye la ubicación exacta, pero puede estar en aproximadamente 0,5 a aproximadamente 15 centímetros, preferentemente en aproximadamente 0,5 a aproximadamente 5 centímetros. El término "separado" como se usa en el presente documento significa que los agentes se administran en un intervalo, por ejemplo, en un intervalo de aproximadamente un día a varias semanas o meses. Los agentes activos pueden administrarse en cualquier orden. El término "secuencialmente" como se usa en el presente documento significa que los agentes se administran en secuencia, por ejemplo, en un intervalo o intervalos de minutos, horas, días o semanas. Si es conveniente los agentes activos pueden administrarse en un ciclo repetitivo regular.

Como se usa en el presente documento, la expresión "secuencia de actuación en *cis*" o "región reguladora en *cis*" o término similar se tomará como que significa cualquier secuencia de nucleótidos que deriva de una secuencia genética expresable en la que la expresión de la secuencia genética se regula, al menos en parte, por la secuencia de nucleótidos. Los expertos en la materia serán conscientes de que una región reguladora en *cis* puede ser capaz de activar, silenciar, potenciar, reprimir o de otra forma alterar el nivel de expresión y/o especificidad de tipo celular y/o especificidad de desarrollo de cualquier secuencia génica estructural.

En toda la presente memoria descriptiva, a menos que el contexto requiera otra cosa, se entenderá que las expresiones "comprenden", "comprende" y "que comprende" implican la inclusión de una etapa indicada o elemento o grupo de etapas o elementos, pero no la exclusión de cualquier otro etapa o elemento o grupo de etapas o elementos.

Como se usa en el presente documento, una "construcción quimérica" se refiere a un polinucleótido que tiene elementos de ácido nucleico heterólogos. Las construcciones quiméricas incluyen "casetes de expresión" o "construcciones de expresión", que se refieren a un conjunto que es capaz de dirigir la expresión de la secuencia o secuencias o gen o genes de interés. Un casete de expresión en general incluye elementos de control tales como un promotor que está unido operativamente al (a fin de dirigir la transcripción de) un polinucleótido sintético y, con frecuencia, también incluye una secuencia de poliadenilación. En ciertas realizaciones de la invención, la construcción quimérica puede estar contenida dentro de un vector. Además de los componentes de la construcción quimérica, el vector puede incluir, uno o más marcadores seleccionables, una señal que permite que el vector exista como ADN monocatenario (por ejemplo, un origen de replicación M13), al menos un sitio de clonación múltiple y un origen "mamífero" de replicación (por ejemplo, un origen de replicación SV40 o de adenovirus).

Por "secuencia codificante" se entiende cualquier secuencia de ácido nucleico que contribuya al código para el producto de polipéptido de un polinucleótido (por ejemplo, un polinucleótido indicador).

Como se usa en el presente documento un "fenotipo conferido" se refiere a un cambio temporal o permanente en el estado de un organismo de interés o clase de organismos de interés, o de una parte o tejido o célula o tipo celular o clase de célula de un organismo de interés, que se produce después de la introducción de un polinucleótido a ese organismo o a esa clase de organismo o a la parte o tejido o célula o tipo celular o clase de célula, o a un precursor de ese organismo o parte o tejido o célula o tipo celular o clase de célula, y que no se habrían producido en ausencia de esta introducción. Normalmente, un cambio temporal o permanente de este tipo se produce como resultado de la transcripción y/o traducción de la información genética contenida dentro de ese polinucleótido en la célula, o en al menos una célula o tipo celular o clase de célula dentro del organismo de interés o dentro de la clase de clase de organismos de interés y puede usarse para distinguir el organismo de interés o clase de organismos de interés, o parte o tejido o célula o tipo celular o clase de célula de los mismos, o progenie genética de éstos, a los que el polinucleótido se les ha proporcionado a partir de un organismo de interés o clase de organismos de interés similares, o parte o tejido o célula o tipo celular o clase de célula de los mismos, o progenie genética de éstos, a los que el polinucleótido no se ha proporcionado.

Como se usa en el presente documento, "respuesta inmunitaria conferida", "respuesta inmunitaria que se confiere" y similares se refieren a un cambio temporal o permanente en la respuesta inmunitaria a un antígeno diana, que se produce o se produciría después de la introducción de un polinucleótido al mamífero y que no se produciría en ausencia de esa introducción. Normalmente, un cambio temporal o permanente de este tipo se produce como resultado de la transcripción y/o traducción de la información genética contenida dentro de ese polinucleótido en una célula, o en al menos una célula o tipo celular o clase de célula en un mamífero o en una clase de mamíferos, y puede usarse para distinguir el mamífero o clase de mamíferos a los que el polinucleótido se les ha proporcionado a partir de un mamífero similar, o clase de mamíferos, a los que no se les ha proporcionado el polinucleótido.

Por "corresponde a" o "correspondiente a" se quiere decir un antígeno que codifica una secuencia de aminoácidos que muestra similitud sustancial a una secuencia de aminoácidos en un antígeno diana. En general, el antígeno mostrará al menos aproximadamente el 30, el 40, el 50, el 55, el 60, el 65, el 70, el 75, el 80, el 85, el 90, el 91, el 92, el 93, el 94, el 95, el 96, el 97, el 98, el 99 % de similitud o de identidad con al menos una parte del antígeno diana (por ejemplo, al menos el 10 %, el 20 %, el 30 %, el 40 %, el 50 %, el 60 %, el 70 %, el 80 %, el 90 % o el 95 % de la secuencia de aminoácidos del antígeno diana).

Por "cantidad eficaz", en el contexto de la modulación de una respuesta inmunitaria o tratar o prevenir una enfermedad o afección, se quiere decir la administración de esa cantidad de la composición a un individuo que lo necesite, ya sea en una dosis única o como parte de una serie, que es eficaz para conseguir esa modulación, tratamiento o prevención. La cantidad eficaz variará dependiendo de la salud y la afección física del individuo que se ha de tratar, el grupo taxonómico del individuo que se ha de tratar, la formulación de la composición, la evaluación de la situación médica y otros factores relevantes. Se espera que la cantidad caiga en un intervalo relativamente amplio que puede determinarse mediante ensayos rutinarios.

Las expresiones "potenciar una respuesta inmunitaria", "producir una respuesta inmunitaria más fuerte" y similares se refieren a aumentar la capacidad de un animal para responder a un antígeno diana (por ejemplo, un antígeno extraño o específico de enfermedad o un autoantígeno), que puede determinarse por ejemplo mediante la detección de un aumento en el número, la actividad y la capacidad de las células del animal que se ceban para atacar dichos antígenos o un aumento en el título o actividad de anticuerpos en el animal, que son inmunointeractivos con el antígeno diana. La potencia de la respuesta inmunitaria puede medirse mediante inmunoensayos convencionales, incluyendo: la medición directa de títulos de anticuerpos o linfocitos de sangre periférica; ensayos de linfocitos T citotóxicos; ensayos de citotoxicidad de linfocitos citotóxicos naturales; ensayos de proliferación celular incluyendo ensayos de linfoproliferación (activación de linfocitos); inmunoensayos de subconjuntos de células inmunitarias;

ensayos de linfocitos T específicos para el antígeno en un sujeto sensibilizado; ensayos cutáneos para la inmunidad mediada por células; etc. Dichos ensayos son bien conocidos en la técnica, véase, por ejemplo, Erickson et al., 1993, *J. Immunol.* 151: 4189-4199; Doe et al., 1994, *Eur. J. Immunol.* 24: 2369-2376. Los métodos recientes de medición de la respuesta inmunitaria mediada por células incluyen la medición de citocinas intracelulares o la secreción de citocinas por poblaciones de linfocitos T o por la medición de linfocitos T específicas de epítipo (por ejemplo, mediante la técnica de tetrámeros) (revisado por McMichael, A. J. y O' Callaghan, C. A., 1998, *J. Exp. Med.* 187 (9) 1367-1371; Mcheyzer-Williams, M. G., et al., 1996, *Immunol. Rev.* 150: 5-21; Lalvani, A., et al., 1997, *J. Exp. Med.* 186: 859-865). Cualquier aumento estadísticamente significativo de la potencia de la respuesta inmunitaria según se mide, por ejemplo, por inmunoensayo se considera una "respuesta inmunitaria potenciada" o "potenciación inmunológica" como se usa en el presente documento. La respuesta inmunitaria potenciada también se indica por manifestaciones físicas tales como fiebre e inflamación, así como por la curación de infecciones sistémicas y locales y la reducción de síntomas en la enfermedad, es decir, la disminución en el tamaño del tumor, el alivio de los síntomas de una enfermedad o afección que incluyen, pero no se limitan a, lepra, tuberculosis, malaria, úlceras aftosas, verrugas herpéticas y papilomatosas, gingivitis, aterosclerosis, acompañantes del SIDA, tales como el sarcoma de Kaposi, infecciones bronquiales y similares. Dichas manifestaciones físicas también abarcan la "respuesta inmunitaria potenciada" o la "potenciación inmunológica" como se usa en el presente documento. Por el contrario, "reducir una respuesta inmunitaria", "producir una respuesta inmunitaria más débil" y similares se refieren a la disminución de la capacidad de un animal para responder a un antígeno diana, que puede determinarse por ejemplo mediante la realización de inmunoensayos o la evaluación de las manifestaciones físicas, como se ha descrito por ejemplo anteriormente.

Los términos "expresión" o "expresión génica" se refieren a la producción de mensajes de ARN y/o a la traducción de mensajes de ARN en proteínas o polipéptidos.

Por "vector de expresión" se entiende cualquier elemento genético autónomo capaz de dirigir la síntesis de una proteína codificada por el vector. Dichos vectores de expresión son conocidos por los profesionales en la materia.

El término "gen" se usa en su contexto más amplio para incluir tanto una región genómica de ADN correspondiente al gen como una secuencia de ADNc correspondiente a exones o una molécula recombinante modificada por ingeniería genética para codificar una forma funcional de un producto.

Como se usa en el presente documento, el término "heterólogo" se refiere a una combinación de elementos que no son de origen natural o que se obtienen de fuentes diferentes.

"Respuesta inmunitaria" o "respuesta inmunológica" se refieren a la acción concertada de linfocitos, células presentadoras de antígeno, células fagocíticas, granulocitos y macromoléculas solubles producidas por las células anteriores o el hígado (incluyendo anticuerpos, citocinas y complemento) que da como resultado el daño selectivo, la destrucción o la eliminación del cuerpo de las células cancerosas, las células tumorales metastásicas, las células de cáncer de mama metastásico, los patógenos invasores, las células o tejidos infectados con patógenos, o, en casos de autoinmunidad o inflamación patológica, las células o tejidos humanos normales. En algunas realizaciones, una "respuesta inmunitaria" abarca el desarrollo en un individuo de una respuesta humoral y/o una respuesta inmunitaria celular a un polipéptido que está codificado por un polinucleótido sintético introducido. Como se sabe en la técnica, las expresiones "respuesta inmunitaria humoral" incluye y abarca una respuesta inmunitaria mediada por moléculas de anticuerpos, mientras que una "respuesta inmunitaria celular" incluye y abarca una respuesta inmunitaria mediada por linfocitos T y/u otros linfocitos de la sangre. Por tanto, una respuesta inmunitaria que es estimulada por un polinucleótido sintético puede ser una que estimule la producción de anticuerpos (por ejemplo, anticuerpos neutralizantes que bloquean toxinas bacterianas y patógenos tales como virus que entran en células y se replican mediante la unión a toxinas y patógenos, normalmente protegiendo a las células de la infección y la destrucción). El polinucleótido sintético también puede desencadenar la producción de linfocitos T citolíticos (LTC) Por tanto, una respuesta inmunológica puede incluir uno o más de los siguientes efectos: la producción de anticuerpos por las células B; y/o la activación de linfocitos T supresoras y/o linfocitos T de memoria/efectoras dirigidas específicamente a un antígeno o antígenos presentes en la composición o vacuna de interés. En algunas realizaciones, estas respuestas pueden servir para neutralizar la infectividad y/o mediar la citotoxicidad anticuerpo-complemento o la citotoxicidad celular dependiente de anticuerpo (CCDA) para proporcionar protección a un hospedador inmunizado. Dichas respuestas pueden determinarse usando inmunoensayos convencionales y ensayos de neutralización, bien conocidos en la técnica. (Véase, por ejemplo, Montefiori et al., 1988, *J Clin Microbiol.* 26: 231-235; Dreyer et al., 1999, *AIDS Res Hum Retroviruses* 15 (17): 1563-1571). El sistema inmunitario innato de los mamíferos también reconoce y responde a las características moleculares de organismos patógenos y células cancerosas a través de la activación de receptores de tipo Toll y moléculas receptoras similares en las células inmunitarias. Tras la activación del sistema inmunitario innato, se activan diversas células de respuesta inmunitaria no adaptativa para, por ejemplo, producir diversas citocinas, linfocinas y quimioquinas. Las células activadas por una respuesta inmunitaria innata incluyen células dendríticas inmaduras y maduras de, por ejemplo, el linaje monocítico y plasmocitoide (MDC, PDC), así como linfocitos T gamma, delta, alfa y beta y células B y similares. Por tanto, la presente divulgación también contempla una respuesta inmunitaria en la que la respuesta inmunitaria implica tanto una respuesta innata como una adaptativa.

Una composición es "inmunógena" si es capaz de ya sea: a) generar una respuesta inmunitaria frente a un antígeno diana (por ejemplo, un antígeno vírico o tumoral) en un individuo; o b) reconstituir, reforzar o mantener una respuesta inmunitaria en un individuo más allá de lo que se produciría si el agente o la composición no se administrara. Un agente o composición son inmunógenos si son capaces de alcanzar cualquiera de estos criterios cuando se administran en dosis únicas o múltiples.

"Inmunomodulación", "modular una respuesta inmunitaria" y similares se refieren a la modulación del sistema inmunitario en respuesta a un estímulo e incluye aumentar o disminuir una respuesta inmunitaria a un antígeno diana o cambiar una respuesta inmunitaria de una que es predominantemente una respuesta inmunitaria humoral a una que es una respuesta inmunitaria más mediada por células y viceversa. Por ejemplo, se sabe en la técnica que la disminución de la cantidad de antígeno para la inmunización puede cambiar el sesgo del sistema inmunitario de una respuesta inmunitaria predominantemente humoral a una respuesta inmunitaria predominantemente celular.

Por "ARN de transferencia isoaceptor" o "iso-ARNt" se quiere decir una o más moléculas de ARN de transferencia que difieren en su secuencia de nucleótidos anticodón, pero son específicos para el mismo aminoácido.

Como se usa en el presente documento, el término "mamífero" se refiere a cualquier mamífero incluyendo, sin limitación, seres humanos y otros primates, incluyendo primates no humanos tales como chimpancés y otros simios y especies de monos; animales de granja tales como vacas, ovejas, cerdos, cabras y caballos; mamíferos domésticos tales como perros y gatos; y animales de laboratorio incluyendo roedores tales como ratones, ratas y conejillos de indias. El término no denota una edad en particular. De este modo, se pretende incluir tanto los individuos adultos como los recién nacidos.

Por "modulación", "modular" y similares se quiere decir aumentar o disminuir, ya sea directa o indirectamente, la cualidad de un fenotipo seleccionado (por ejemplo, una respuesta inmunitaria). En ciertas realizaciones, "modulación" o "modular" significa que una respuesta inmunitaria deseada/seleccionada es más eficiente (por ejemplo, al menos un 10 %, un 20 %, un 30 %, un 40 %, un 50 %, un 60 % o más), más rápida (por ejemplo, al menos un 10 %, un 20 %, un 30 %, un 40 %, un 50 %, un 60 % o más), mayor en magnitud (por ejemplo, al menos un 10 %, un 20 %, un 30 %, un 40 %, un 50 %, un 60 % o más), y/o más fácilmente inducible (por ejemplo, al menos un 10 %, un 20 %, un 30 %, un 40 %, un 50 %, un 60 % o más) que si se hubiera utilizado el polinucleótido parental en las mismas condiciones que el polinucleótido sintético. En otras realizaciones, "modulación" o "modular" significa cambiar una respuesta inmunitaria de una respuesta inmunitaria predominantemente mediada por anticuerpos como es conferida por el polinucleótido parental, a una respuesta inmunitaria predominantemente celular como es conferida por el polinucleótido sintético en las mismas condiciones. En otras realizaciones más, "modulación" o "modular" significa cambiar una respuesta inmunitaria de una respuesta inmunitaria predominantemente celular como es conferida por el polinucleótido parental, a una respuesta inmunitaria predominantemente mediada por anticuerpos como es conferida por el polinucleótido sintético en las mismas condiciones.

Por "gen natural" se quiere decir un gen que codifica naturalmente la proteína. Sin embargo, es posible que el polinucleótido parental codifique una proteína que no sea de origen natural pero que se ha diseñado usando técnicas recombinantes.

La expresión "región 5' no codificante" se usa en el presente documento en su contexto más amplio para incluir todas las secuencias de nucleótidos que derivan de la región aguas arriba de un gen expresable, distintas de aquellas secuencias que codifican restos de aminoácido que comprenden el producto polipeptídico del gen, en el que la región 5' no codificante confiere o activa o facilita de otro modo, al menos en parte, la expresión del gen.

El término "oligonucleótido" como se usa en el presente documento se refiere a un polímero compuesto de una multiplicidad de unidades de nucleótidos (desoxirribonucleótidos o ribonucleótidos, o variantes estructurales relacionadas o análogos sintéticos de los mismos) unidos a través de enlaces fosfodiéster (o variantes estructurales relacionadas o análogos sintéticos de los mismos). Por tanto, aunque el término "oligonucleótido" normalmente se refiere a un polímero de nucleótidos en el que los nucleótidos y los enlaces entre ellos son de origen natural, se entenderá que el término también incluye dentro de su alcance diversos análogos que incluyen, pero no se limitan a, ácidos nucleicos peptídicos (ANP), fosforamidatos, fosforotioatos, fosfonatos de metilo, ácidos 2-O-metil ribonucleicos y similares. El tamaño exacto de la molécula puede variar dependiendo de la aplicación particular. Un oligonucleótido normalmente es bastante corto de longitud, generalmente de aproximadamente 10 a 30 nucleótidos, pero el término puede referirse a moléculas de cualquier longitud, aunque el término "polinucleótido" o "ácido nucleico" se usa normalmente para oligonucleótidos grandes.

Las expresiones "conectado operativamente", "unido operativamente" y similares, como se usan en el presente documento, se refieren a una disposición de elementos en la que los componentes descritos de este modo se configuran de modo que realizan su función habitual. Por tanto, un promotor dado operativamente unido a una secuencia codificante es capaz de efectuar la expresión de la secuencia codificante cuando están presentes las enzimas apropiadas. El promotor no necesita estar contiguo a la secuencia codificante, siempre que funcione para dirigir la expresión del mismo. De este modo, por ejemplo, puede haber presentes secuencias intermedias no traducidas pero transcritas entre la secuencia promotora y la secuencia codificante y la secuencia promotora todavía

puede considerarse "operativamente unida" a la secuencia codificante. Expresiones tales como "conectado operativamente", por tanto, incluyen poner un gen estructural bajo el control regulador de un promotor, que después controla la transcripción y opcionalmente la traducción del gen. En la construcción de combinaciones de genes promotores/estructurales heterólogos, en general se prefiere poner la secuencia genética o promotor a una distancia desde el sitio de inicio de la transcripción génica que es aproximadamente la misma que la distancia entre esa secuencia genética o promotor y el gen que controla en su entorno natural; es decir, el gen del que deriva la secuencia genética o promotor. Como se sabe en la técnica, puede acomodarse alguna variación en esta distancia sin pérdida de función. De forma similar, el posicionamiento preferido de un elemento de secuencia reguladora con respecto a un gen heterólogo que se ha de colocar bajo su control se define por el posicionamiento del elemento en su entorno natural; es decir, los genes de los que deriva.

Por "vehículo farmacéuticamente aceptable" se quiere decir una carga, un diluyente o una sustancia encapsulante sólidos o líquidos que pueden usarse con seguridad en la administración tópica o sistémica.

El término "fenotipo" significa una cualquiera o más características, propiedades, atributos o rasgos físicos o funcionales detectables de un organismo, tejido o célula o clase de organismos, tejidos o células, que en general dan como resultado la interacción entre la composición genética (es decir, el genotipo) del organismo, tejido o célula o clase de organismos, tejidos o células y el entorno. En ciertas realizaciones, el término "fenotipo" excluye la resistencia a un agente selectivo o la detección de una actividad enzimática o emisora de luz, conferidas directamente por una proteína indicadora.

Por "preferencia fenotípica" se quiere decir la preferencia con la que un organismo usa un codón para producir un fenotipo seleccionado. Esta preferencia puede evidenciarse, por ejemplo, por la cualidad de un fenotipo seleccionado que es producible por un polinucleótido que comprende el codón en un marco abierto de lectura que codifica un polipéptido que produce el fenotipo seleccionado. En cierta realización, la preferencia de uso es independiente de la vía por la que se introduce el polinucleótido en el organismo. Sin embargo, en otras realizaciones, la preferencia de uso depende de la vía de introducción del polinucleótido en el organismo.

Las expresiones "polinucleótido" o "ácido nucleico" como se usan en el presente documento designan ARNm, ARN, ARNc, ADNc o ADN. El término se refiere normalmente a oligonucleótidos mayores de 30 nucleótidos de longitud.

"Polipéptido", "péptido" y "proteína" se usan indistintamente en el presente documento para referirse a un polímero de restos de aminoácidos y a variantes y análogos sintéticos del mismo. Por tanto, estos términos se aplican a polímeros de aminoácidos en los que uno o más restos de aminoácidos son un aminoácido de origen no natural sintético, tal como un análogo químico de un aminoácido de origen natural correspondiente, así como a polímeros de aminoácidos de origen natural. Como se usa en el presente documento, los términos "polipéptido", "péptido" y "proteína" no se limitan a una longitud mínima del producto. De este modo, los péptidos, oligopéptidos, dímeros, multímeros y similares, se incluyen dentro de la definición. Tanto las proteínas de longitud completa como los fragmentos de las mismas están incluidos en la definición. Los términos también incluyen modificaciones posteriores a la expresión de un polipéptido, por ejemplo, glucosilación, acetilación, fosforilación y similares. En algunas realizaciones, un "polipéptido" se refiere a una proteína que incluye modificaciones, tales como delecciones, adiciones y sustituciones (en general de naturaleza conservadora) de la secuencia nativa, a condición de que la proteína mantenga la actividad deseada. Estas modificaciones pueden ser deliberadas, como a través de mutagénesis dirigida, o pueden ser accidentales, tales como a través de mutaciones de hospedadores que producen las proteínas o errores debidos a la amplificación por PCR.

Las expresiones "variante de polipéptido" y "variante" se refieren a polipéptidos que varían con respecto a un polipéptido de referencia mediante la adición, delección o sustitución (en general de naturaleza conservadora) de al menos un resto de aminoácido. Normalmente, las variantes conservan una actividad deseada del polipéptido de referencia, tal como la actividad antigénica en la inducción de una respuesta inmunitaria frente a un antígeno diana. En general, los polipéptidos variantes son "sustancialmente similares" o "sustancialmente idénticos" al polipéptido de referencia, por ejemplo, una identidad o similitud de secuencia de aminoácidos de más del 50 %, en general de más del 60 %-70 %, incluso más en particular del 80 %-85 % o más, tal como al menos el 90 %-95 % o más, cuando las dos secuencias están alineadas. Con frecuencia, las variantes incluirán el mismo número de aminoácidos, pero también incluirán sustituciones, como se explica en el presente documento.

Las expresiones "célula o tejido precursores" y "célula o tejido progenitores" como se usa en el presente documento se refieren a una célula o un tejido que pueden dar lugar a una célula o un tejido particulares en los que se produce un polipéptido mediante la expresión de las secuencias codificantes en las construcciones sintéticas de la invención.

Los términos "precursor" y "progenitor", como se usan en el presente documento en el contexto de la preferencia fenotípica, se refieren a una célula o una parte de organismo que pueden dar lugar a un organismo de interés en los que se desea la expresión fenotípica o en los que se ha de determinar la preferencia fenotípica de un codón.

Por "cebador" se quiere decir un oligonucleótido que, cuando se aparea con una cadena de ADN, es capaz de iniciar la síntesis de un producto de prolongación del cebador en presencia de un agente de polimerización adecuado. El

cebador es preferentemente monocatenario para una máxima eficiencia en la amplificación, pero como alternativa puede ser bicatenario. Un cebador debe ser lo suficientemente largo para cebar la síntesis de productos de prolongación en presencia del agente de polimerización. La longitud del cebador depende de muchos factores, incluyendo la aplicación, la temperatura que se ha de emplear, las condiciones de reacción del molde, otros reactivos y la fuente de cebadores. Por ejemplo, dependiendo de la complejidad de la secuencia diana, el cebador oligonucleotídico contiene normalmente de 15 a 35 o más nucleótidos, aunque puede contener menos nucleótidos. Los cebadores pueden ser polinucleótidos grandes, tales como de aproximadamente 200 nucleótidos a varias kilobases o más. Los cebadores pueden seleccionarse para que sean "sustancialmente complementarios" a la secuencia en el molde diseñado para que se hibridara con ellos y para servir como un sitio para la iniciación de la síntesis. Por "sustancialmente complementario", se quiere decir que el cebador es suficientemente complementario para hibridarse con una secuencia de nucleótidos diana. Preferentemente, el cebador no contiene ningún desapareamiento con el molde diseñado para que se hibridara con él, pero esto no es esencial. Por ejemplo, los nucleótidos no complementarios pueden estar unidos al extremo 5' del cebador, siendo el resto de la secuencia del cebador complementario al molde. Como alternativa, pueden intercarse nucleótidos no complementarios o un tramo de nucleótidos no complementarios en un cebador, a condición de que la secuencia del cebador tenga suficiente complementariedad con la secuencia del molde que se ha de hibridar con ella y, de ese modo, forme un molde para la síntesis del producto de prolongación del cebador.

Por "producir" y términos similares tales como "producción" y "producible", en el contexto de la producción de proteínas, se quiere decir la producción de una proteína en un nivel suficiente para conseguir una función o fenotipo particulares asociados a la proteína. Por el contrario, las expresiones "no producible" y "sustancialmente no producible" como se usan indistintamente en el presente documento se refieren a (a) la no producción de una proteína, (b) la producción de una proteína en un nivel que no es suficiente para efectuar una función o fenotipo particulares asociados a la proteína, (c) la producción de una proteína, que no puede ser detectado por un anticuerpo monoclonal específico para la proteína o (d) la producción de una proteína, que es inferior al 1 % del nivel producido en una celular de tipo silvestre que normalmente produce la proteína.

La referencia en el presente documento a un "promotor" se ha de tomar en su contexto más amplio e incluye las secuencias reguladoras transcripcionales de un gen genómico clásico, incluyendo la caja TATA que es necesaria para el inicio preciso de la transcripción, con o sin una secuencia de caja CCAAT y elementos reguladores adicionales (es decir, secuencias activadoras aguas arriba, potenciadores y silenciadores) que alteran la expresión génica en respuesta a estímulos de desarrollo y/o ambientales o de una forma específica de tejido o específica de tipo celular. Un promotor por lo general, pero no necesariamente, se sitúa aguas arriba o en posición 5', de un gen estructural, cuya expresión regula. Además, los elementos reguladores que comprenden un promotor por lo general se sitúan a 2 kb del sitio de inicio de la transcripción del gen. Los promotores preferidos de acuerdo con la invención pueden contener copias adicionales de uno o más elementos reguladores específicos para potenciar adicionalmente la expresión en una célula y/o para alterar el momento de la expresión de un gen estructural al que está conectado operativamente.

El término "cualidad" se usa en el presente documento en su sentido más amplio e incluye una medición, potencia, intensidad, nivel o grado de un fenotipo, por ejemplo, una respuesta inmunitaria superior o inferior, una resistencia a la enfermedad aumentada o disminuida, una acumulación de sacarosa superior o inferior, una tolerancia a la sal mejor o peor etc.

Por "elemento regulador" o "secuencia reguladora" se quiere decir una secuencia de ácido nucleico (por ejemplo, ADN) que expresa una secuencia de nucleótidos unida operativamente (por ejemplo, una secuencia codificante) en una célula hospedadora particular. Las secuencias reguladoras que son adecuadas para células procariotas, por ejemplo, incluyen un promotor y, opcionalmente, una secuencia de actuación *cis* tal como una secuencia operadora y un sitio de unión a ribosomas. Las secuencias de control que son adecuadas para las células eucariotas incluyen promotores, señales de poliadenilación, potenciadores de la transcripción, potenciadores de la traducción, secuencias líderes o traseras que modulan la estabilidad del ARNm, así como secuencias de dirección que se dirigen un producto codificado por un polinucleótido transcrito a un compartimiento intracelular dentro de una célula o al medio extracelular.

La expresión "identidad de secuencia" como se usa en el presente documento se refiere al grado en que las secuencias son idénticas nucleótido a nucleótido o aminoácido a aminoácido en una ventana de comparación. Por tanto, un "porcentaje de identidad de secuencia" se calcula comparando dos secuencias óptimamente alineadas en la ventana de comparación, determinando el número de posiciones en las que la base de ácido nucleico idéntica (por ejemplo, A, T, C, G, I) o el resto de aminoácido idéntico (por ejemplo, Ala, Pro, Ser, Thr, Gly, Val, Leu, Ile, Phe, Tyr, Trp, Lys, Arg, His, Asp, Glu, Asn, Gln, Cys y Met) que se producen en ambas secuencias para proporcionar el número de posiciones coincidentes, dividiendo el número de posiciones coincidentes por el número total de posiciones en la ventana de comparación (es decir, el tamaño de la ventana) y multiplicando el resultado por 100 para proporcionar el porcentaje de identidad de secuencia. Para los fines de la presente invención, se entenderá que "identidad de secuencia" significa el "porcentaje de coincidencia" calculado mediante el programa de ordenador DNASIS (Versión 2.5 para Windows, disponible de Hitachi Software Engineering Co., Ltd., San Francisco Sur,

California, EE.UU.), usando valores predeterminados convencionales tales como se usan en el manual de referencia que acompaña al software.

"Similitud" se refiere al número en porcentaje de aminoácidos que son idénticos o constituyen sustituciones conservadoras como se definen en la Tabla 10. La similitud puede determinarse usando programas de comparación de secuencias, tales como GAP (Deveraux et al. 1984, *Nucleic Acids Research* 12, 387-395). De esta manera, las secuencias de una longitud similar o sustancialmente diferente de las citadas en el presente documento pueden compararse por inserción de huecos en la alineación, determinándose dichos huecos, por ejemplo, mediante el algoritmo de comparación utilizado por GAP.

Los términos utilizados para describir relaciones de secuencia entre dos o más polinucleótidos o polipéptidos incluyen "secuencia de referencia", "ventana de comparación", "identidad de secuencia", "porcentaje de identidad de secuencia" e "identidad sustancial". Una "secuencia de referencia" tiene al menos 12 pero frecuentemente de 15 a 18 y con frecuencia al menos 25 unidades de monómero, incluyendo los nucleótidos y restos de aminoácidos, de longitud. Debido a que dos polinucleótidos pueden comprender cada uno (1) una secuencia (es decir, sólo una porción de la secuencia de polinucleótidos completa) que es similar entre los dos polinucleótidos y (2) una secuencia que es divergente entre los dos polinucleótidos, las comparaciones de secuencias entre dos (o más) polinucleótidos se realizan normalmente comparando secuencias de los dos polinucleótidos sobre una "ventana de comparación" para identificar y comparar regiones locales de similitud de secuencia. Una "ventana de comparación" se refiere a un segmento conceptual de al menos 6 posiciones contiguas, por lo general de aproximadamente 50 a aproximadamente 100, más por lo general de aproximadamente 100 a aproximadamente 150 en el que una secuencia se compara con una secuencia de referencia del mismo número de posiciones contiguas después de que la dos secuencias se alinean de manera óptima. La ventana de comparación puede comprender adiciones o deleciones (es decir, huecos) de aproximadamente el 20 % o menos en comparación con la secuencia de referencia (que no comprende adiciones o deleciones) para la alineación óptima de las dos secuencias. La alineación óptima de secuencias para alinear una ventana de comparación puede realizarse mediante implementaciones computarizadas de algoritmos (GAP, BESTFIT, FASTA y TFASTA en el Paquete de Software Wisconsin Genetics versión 7.0, Genetics Computer Group, 575 Science Drive Madison, WI, EE.UU.) o mediante la inspección y la mejor alineación (es decir, que da como resultado el mayor porcentaje de homología en la ventana de comparación) generada mediante cualquiera de los diversos métodos seleccionados. También puede hacerse referencia a la familia de programas BLAST como se desvela, por ejemplo, por Altschul et al., 1997, *Nucl. Acids Res.* 25: 3389. Pueden encontrarse un análisis detallado de análisis de la secuencia en la Unidad 19.3 de Ausubel et al., *"Current Protocols in Molecular Biology"*, John Wiley & Sons Inc, 1994-1998, Capítulo 15.

Como se usa en el presente documento, la expresión "par de unión específica" se refiere a un par de moléculas que interactúan físicamente entre sí de una manera específica que da lugar a una actividad biológica, es decir, a la exclusión sustancial de otros polipéptidos. Los miembros de un par de unión específica interactúan a través de dominios de interacción complementarios, de manera que interactúan con la exclusión sustancial de proteínas que no tienen un dominio de interacción complementario. Los ejemplos no limitantes de pares de unión específica incluyen pares anticuerpo-antígeno, pares enzima-sustrato, factores de transcripción dimericos (por ejemplo, AP-1, compuesto de Fos específicamente unido a Jun a través de un dominio de interacción de cremallera de leucina) y pares receptor-ligando.

Las expresiones "polinucleótido sintético", "construcción sintética" y similares, como se usan en el presente documento, se refieren a una molécula de ácido nucleico que se forma mediante técnicas recombinantes o de síntesis y normalmente incluye polinucleótidos que no se encuentran normalmente en la naturaleza.

La expresión "codón sinónimo" como se usa en el presente documento se refiere a un codón que tiene una secuencia de nucleótidos diferente de otro codón pero que codifica el mismo aminoácido que ese otro codón.

Por "tratamiento", "tratar", "tratado" y similares se pretende incluir tanto el tratamiento terapéutico como el profiláctico.

Por "vector" se entiende una molécula de ácido nucleico, preferentemente una molécula de ADN derivada, por ejemplo, de un plásmido, bacteriófago o virus de planta, en la que puede insertarse o clonarse una secuencia de ácido nucleico. Un vector contiene preferentemente uno o más sitios de restricción únicos y puede ser susceptible de replicación autónoma en una célula hospedadora definida incluyendo una célula o tejido diana o una célula o tejido progenitores de los mismos, o puede integrarse con el genoma del hospedador definido de manera que la secuencia clonada sea reproducible. En consecuencia, el vector puede ser un vector de replicación autónoma, es decir, un vector que existe como una entidad extracromosómica, cuya replicación es independiente de la replicación cromosómica, por ejemplo, un plásmido circular lineal o cerrado, un elemento extracromosómico, un minicromosoma o un cromosoma artificial. El vector puede contener cualquier medio para garantizar la autorreplicación. Como alternativa, el vector puede ser uno que, cuando se introduce en la célula hospedadora, se integra en el genoma y se replica junto con el cromosoma o cromosomas en los que se ha integrado. Un sistema de vectores puede comprender un único vector o plásmido, dos o más vectores o plásmidos, que juntos contienen el ADN total que se ha de introducir en el genoma de la célula hospedadora, o un transposón. La elección del vector normalmente

dependerá de la compatibilidad del vector con la célula hospedadora en la que el vector se ha de introducir. El vector también puede incluir un marcador de selección tal como un gen de resistencia a antibióticos que puede usarse para la selección de transformantes adecuados. Los ejemplos de dichos genes de resistencia son bien conocidos por los expertos en la materia.

2. Abreviaturas

Se usan las siguientes abreviaturas en toda la solicitud:

nt = nucleótido
nts = nucleótidos
aa = aminoácido o aminoácidos
kb = kilobase o kilobases o par o pares de kilobases
kDa = kilodalton o kilodaltons
d = día
h = hora
s = segundo

3. Sistema de construcciones de la invención

De acuerdo con la presente invención, se proporciona un sistema de construcciones para determinar la eficiencia de traducción o la preferencia fenotípica de diferentes codones sinónimos. En su forma más amplia, el sistema comprende una pluralidad de construcciones sintéticas cada una de las cuales es útil para interrogar la eficiencia de traducción o la preferencia fenotípica de un solo codón ("codón de interrogación"), en el que el codón de interrogación de una construcción es diferente del codón de interrogación de otra. Por tanto, con el fin de comparar la eficiencia de traducción o la preferencia fenotípica de diferentes codones sinónimos, en general es deseable usar dos o más construcciones sintéticas, convenientemente una para cada codón sinónimo que codifica un aminoácido particular. Por ejemplo, en el caso de la arginina, son necesarias 6 construcciones sintéticas para determinar la eficiencia de traducción o la preferencia fenotípica de los 6 codones sinónimos para la arginina (es decir, Arg^{CGA}, Arg^{CGC}, Arg^{CGT}, Arg^{AGA}, Arg^{AGG}, Arg^{CGG}). Por el contrario, solo se necesitan 2 construcciones sintéticas para determinar la eficiencia de traducción o la preferencia fenotípica de ambos codones sinónimos para la fenilalanina (es decir, Phe^{TTT}, Phe^{TTC}) y así sucesivamente. En consecuencia, con el fin de interrogar la eficiencia de traducción o la preferencia fenotípica de un número finito de codones sinónimos, en general se necesitará un número correspondiente de construcciones sintéticas.

Las construcciones sintéticas de la invención comprenden, cada una, una secuencia reguladora que está conectada operativamente a un polinucleótido indicador, en las que el polinucleótido indicador de una construcción respectiva codifica la misma secuencia de aminoácidos que el polinucleótido indicador de otra. De acuerdo con la presente invención, los polinucleótidos indicadores individuales usan el mismo codón de interrogación para codificar un aminoácido particular en una o más posiciones de la secuencia de aminoácidos, en los que el codón de interrogación de un polinucleótido indicador es diferente, pero sinónimo, del codón de interrogación de otro. En realizaciones específicas, las secuencias codificantes de polinucleótidos indicadores individuales comprenden el mismo número de codones de interrogación. Convenientemente, todos los codones en una secuencia codificante respectiva, que codifican un aminoácido particular, son el mismo codón de interrogación. Sin embargo, esto no es necesario, ya que es posible usar un menor número de codones de interrogación que el número de codones en una secuencia codificante respectiva, que codifican el mismo aminoácido que los codones de interrogación. Sin embargo, la sensibilidad de una construcción sintética individual para determinar la eficiencia de traducción o la preferencia fenotípica de un codón de interrogación correspondiente se mejora en general mediante la incorporación de más codones de interrogación en la secuencia codificante.

En algunas realizaciones, el codón o codones de interrogación en una secuencia codificantes se sitúan en las mismas posiciones que los codones de interrogación en otra secuencia codificante. En otras realizaciones, el codón o codones de interrogación de una secuencia codificante se sitúan en diferentes posiciones con respecto a los codones de interrogación en otra secuencia codificante. Por ejemplo, una primera secuencia codificante y una segunda secuencia codificante pueden contener, cada una, 5 codones que codifican un aminoácido particular y solo 3 de ellos se usan como codón de interrogación. En este ejemplo no limitante, la primera secuencia codificante puede comprender la secuencia:

X₁ X₂ X₃ A₁ X₄ X₅ B₁ X₆ X₇ X₈ A₂ X₉ A₃ X₁₀ X₁₁ X₁₂ B₂ X₁₃ X₁₄

y la segunda secuencia codificante puede comprender:

X₁ X₂ X₃ A₁ X₄ X₅ A₂ X₆ X₇ X₈ B₁ X₉ A₃ X₁₀ X₁₁ X₁₂ B₂ X₁₃ X₁₄

en las que:

A₁₋₃ representan el mismo codón de interrogación;

B₁₋₂ representan codones que codifican el mismo aminoácido que el codón de interrogación; y

5 X₁₋₁₄ representan codones que codifican aminoácidos diferentes que el aminoácido codificado por A₁₋₃ y B₁₋₂;

En algunas realizaciones, el sistema de construcciones comprende construcciones sintéticas para interrogar la eficiencia de traducción o la preferencia fenotípica de codones que codifican dos o más aminoácidos diferentes. En ejemplos ilustrativos de este tipo, el sistema de construcciones comprende construcciones sintéticas para interrogar la eficiencia de traducción o la preferencia fenotípica de codones que codifican 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 o 20 aminoácidos (convenientemente de origen natural). En realizaciones específicas, el sistema de construcciones comprende 59 construcciones sintéticas para interrogar la eficiencia de traducción o la preferencia fenotípica de todos los codones de origen natural para los que existen dos o más codones sinónimos (por ejemplo, Ala^{GCT}, Ala^{GCC}, Ala^{GCA}, Ala^{GCG}, Arg^{CGA}, Arg^{CGC}, Arg^{CGT}, Arg^{AGA}, Arg^{AGG}, Arg^{CGG}, Asn^{AAC}, Asn^{AAT}, Asp^{GAC}, Asp^{GAT}, Cys^{TGC}, Cys^{TGT}, Glu^{GAA}, Glu^{GAG}, Gln^{CAA}, Gln^{CAG}, Gly^{GGA}, Gly^{GGG}, Gly^{GGT}, Gly^{GGC}, His^{CAC}, His^{CAT}, Ile^{ATC}, Ile^{ATT}, Ile^{ATA}, Leu^{CTG}, Leu^{CTC}, Leu^{CTA}, Leu^{CTT}, Leu^{TTG}, Leu^{TTA}, Lys^{AAG}, Lys^{AAA}, Phe^{TTT}, Phe^{TTT}, Pro^{CCC}, Pro^{CCT}, Pro^{CCA}, Pro^{CCG}, Ser^{TCC}, Ser^{TCT}, Ser^{TCA}, Ser^{TCC}, Ser^{AGC}, Ser^{AGT}, Thr^{ACC}, Thr^{ACA}, Thr^{ACT}, Tyr^{TAC}, Tyr^{TAT}, Val^{GTG}, Val^{GTC}, Val^{GTT} y Val^{GTA}).

En algunas realizaciones en las que el sistema de construcciones se usa para determinar la eficiencia de traducción de codones sinónimos, el polinucleótido indicador codifica una secuencia de aminoácidos que define, en su totalidad o en parte, una proteína indicadora que, cuando está presente en una célula, es detectable y distinguible de otros polipéptidos presentes en la célula. Una proteína indicadora puede ser una proteína de origen natural o una proteína que no es de origen natural. Los ejemplos ilustrativos de dichas proteínas indicadoras incluyen proteínas fluorescentes tales como la proteína verde fluorescente (gfp), la proteína cian fluorescente (cfp), la proteína roja fluorescente (prf) o la proteína azul fluorescente (bfp) o derivados de estas proteínas, o proteínas enzimáticas tales como cloranfenicol acetil transferasa, β -galactosidasa, β -glucuronidasa (GUS), fosfatasa alcalina placentaria secretada y β -lactamasa, proteínas quimioluminiscentes tales como luciferasa y proteínas marcadoras seleccionables incluyendo proteínas codificadas por genes de resistencia a antibióticos (por ejemplo, genes de resistencia a higromicina, genes de resistencia a neomicina, genes de resistencia a tetraciclina, genes de resistencia a ampicilina, genes de resistencia a kanamicina, genes de resistencia a fleomicina, genes de resistencia a herbicidas tales como el gen de resistencia a bialafos (BAR) que confieren resistencia al herbicida BASTA, genes de resistencia a bleomicina, genes de resistencia a genéticina, genes de resistencia a carbenicilina, genes de resistencia a cloranfenicol, genes de resistencia a puomicina, genes de blasticidina-S-desaminasa), genes de resistencia a metales pesados, genes *hisD*, genes de hipoxantina fosforribosil transferasa (HPRT) y genes de guanina fosforribosil transferasa (Gpt).

En algunas realizaciones en las que el sistema de construcciones se usa para determinar la preferencia fenotípica de codones sinónimos, el polinucleótido indicador codifica una secuencia de aminoácidos que define, en su totalidad o en parte, que una proteína indicadora confiere a un organismo de interés o parte del mismo, ya sea por sí misma o en asociación con otras moléculas, un fenotipo seleccionado o un fenotipo de la misma clase que el fenotipo seleccionado. Por ejemplo, la proteína indicadora puede ser un polipéptido asociado a fenotipo (por ejemplo, un antígeno específico de melanoma tal como BAGE o GAGE-1) que será el objeto de producción del fenotipo seleccionado (por ejemplo, inmunidad a melanoma). Como alternativa, el polipéptido asociado a fenotipo (por ejemplo, proteína fluorescente verde o un antígeno asociado al sistema gastrointestinal tal como 17-1A) puede no producir el fenotipo seleccionado (por ejemplo, inmunidad a melanoma) pero puede producir la misma clase de fenotipo (por ejemplo, una respuesta inmunitaria) que el fenotipo seleccionado. En ejemplos ilustrativos, el polipéptido asociado a fenotipo se selecciona entre antígenos que incluyen antígenos de organismos patógenos o cánceres (por ejemplo, en los que el fenotipo es la inmunidad a la enfermedad) y autoantígenos o antígenos de trasplante (por ejemplo, en los que el fenotipo es la anergia o tolerancia específica de antígeno), factores de crecimiento (por ejemplo, en los que el fenotipo se selecciona entre el tamaño del organismo o parte del mismo, la cicatrización de heridas, la proliferación celular, la diferenciación celular, la migración celular, la función de células inmunitarias), hormonas (por ejemplo, en las que el fenotipo se la lactancia aumentada, por ejemplo, usando oxitocina, o el alivio de un estado diabético, por ejemplo, usando insulina) y toxinas (por ejemplo, en las que el fenotipo es la regresión tumoral o la muerte celular). En realizaciones específicas, el fenotipo o clase de fenotipo seleccionado corresponde a un estado o una condición beneficiosos o mejorados o superiores del organismo o parte del mismo con respecto a un estado o estado de referencia. En ejemplos ilustrativos, el estado o condición de referencia corresponde a un estado patofisiológico. Los fenotipos contemplados por la presente invención incluyen cualquier rasgo beneficioso deseable incluyendo, pero no limitado a: la inmunidad (por ejemplo, la inmunidad a la infección por patógenos o al cáncer); la tolerancia a antígenos (por ejemplo, la anergia de linfocitos T específica de antígeno, la tolerancia a alérgenos, los antígenos de trasplante y los autoantígenos); la angiogénesis (por ejemplo, la formación de vasos sanguíneos en el corazón y vasculatura y en crecimientos tumorales); la anti-angiogénesis (por ejemplo, el tratamiento de la cardiopatía isquémica y de tumores); la mejora de síntomas clínicos (por ejemplo, fiebre, inflamación; encefalitis; pérdida de peso; anemia; síntomas sensitivos, tales como parestesia o hipoestesia; ataxia; neuralgia; parálisis; vértigo; anomalías urinarias o del movimiento intestinal; y disfunción cognitiva tal como pérdida de la memoria, atención alterada, dificultades para resolver problemas, procesamiento ralentizado de

la información y dificultad para cambiar entre tareas cognitivas); la muerte celular reducida o aumentada (por ejemplo, apoptosis); la diferenciación celular reducida o aumentada; la proliferación celular reducida o aumentada; la regresión del tumor o cáncer; el crecimiento y reparación del tejido u órgano; la disminución de la fibrosis; la inhibición o la reversión de la senescencia celular; el aumento o la reducción de la migración celular; la expresión diferencial de proteínas entre las diferentes células o tejidos de un organismo o parte del mismo; la recuperación de traumatismos; la recuperación de quemaduras; la resistencia o la sensibilidad a antibióticos (por ejemplo, la resistencia o la sensibilidad a antibióticos aminoglucósidos tales como geneticina y paromomicina); la tolerancia o la sensibilidad a herbicidas (por ejemplo, la tolerancia o la sensibilidad al glifosato o al glufosinato); la biosíntesis o la modificación de almidón (por ejemplo, usando una enzima ramificadora de almidón, sintasas de almidón, ADP-glucosa pirofosforilasa); la biosíntesis de ácidos grasos (por ejemplo, usando una desaturasa o hidroxilasa); la resistencia o la tolerancia a enfermedades (por ejemplo, la resistencia a enfermedades animales tales como enfermedad cardiovascular, autoinmunidad, enfermedad de Alzheimer, enfermedad de Parkinson, diabetes, SIDA etc o la resistencia a enfermedades de plantas tales como roya, enanismo, putrefacción, carbón, moho, sarna y mildiú); la resistencia o la tolerancia a las plagas incluyendo la resistencia o la tolerancia a los insectos (por ejemplo, la resistencia a la polilla y a gusanos); la resistencia o la tolerancia a virus (por ejemplo, la resistencia a virus animales tales como herpesvirus, hepadnavirus, adenovirus, flavivirus, lentivirus, poxvirus etc o la resistencia a virus de plantas tales como badnavirus, caulimovirus, potivirus, luteovirus, rhabdovirus etc); la resistencia o la tolerancia a hongos (por ejemplo, la resistencia a hongos micorrízicos arbusculares, hongos endófitos etc); un rasgo metabólico incluyendo el metabolismo de la sacarosa (por ejemplo, la isomerización de sacarosa); la resistencia o la tolerancia a las heladas; la tolerancia al estrés (por ejemplo, la tolerancia a la sal, la tolerancia a la sequía); y la mejora del contenido de los alimentos o el aumento de los rendimientos. Los expertos en la materia reconocerán que las clases de fenotipo ejemplares anteriores pueden subdividirse en subclases fenotípicas y que dichas subclases también pertenecerían al alcance de las clases fenotípicas contempladas por la presente invención. Por ejemplo, las subclases de inmunidad incluyen la inmunidad innata (que puede subdividirse adicionalmente, entre otros, en sistema del complemento, monocitos, macrófagos, neutrófilos y linfocitos citolíticos naturales), la inmunidad celular (que puede subdividirse adicionalmente, entre otros, en linfocitos T citolíticos, células dendríticas y linfocitos T auxiliares) y la inmunidad humoral (que puede subdividirse adicionalmente, entre otros, en las subclases de anticuerpos IgA, IgD, IgE, IgG e IgM).

En algunas realizaciones, el polinucleótido indicador de construcciones sintéticas individuales comprende adicionalmente una secuencia codificante auxiliar que codifica un marcador detectable (por ejemplo, estreptavidina, avidina, un anticuerpo, un antígeno, un epítipo, un hapteno, una proteína o un resto fluorescente, quimioluminiscente o químicamente reactivo). El marcador detectable es convenientemente un miembro de un par de unión específica, que incluye, por ejemplo, pares anticuerpo-antígeno (o hapteno), pares ligando-receptor, pares enzima-sustrato, pares biotina-avidina y similares. En ejemplos ilustrativos de este tipo, la secuencia codificante auxiliar de un polinucleótido indicador codifica un primer marcador (por ejemplo, un primer epítipo al que se une un primer anticuerpo) y la secuencia codificante auxiliar de otro polinucleótido indicador codifica un segundo marcador (por ejemplo, un segundo epítipo al que se une un segundo anticuerpo), que es distinguible de forma detectable del primer marcador. En estos ejemplos, es posible distinguir de forma detectable los productos polipeptídicos de diferentes polinucleótidos indicadores en la misma célula u organismo de interés o parte de los mismos, permitiendo de este modo la determinación simultánea de las eficiencias de traducción o las preferencias fenotípicas de diferentes codones de interrogación en la misma célula u organismo o parte de los mismos.

De acuerdo con la presente invención, el polinucleótido indicador está unido operativamente en las construcciones sintéticas a una secuencia reguladora. La secuencia reguladora comprende convenientemente secuencias control de la transcripción y/o de la traducción, que serán compatibles para la expresión en la célula u organismo de interés. Normalmente, las secuencias de control reguladoras de la transcripción y de la traducción incluyen, pero no se limitan a, una secuencia promotora, una región 5' no codificante, una región reguladora en *cis* tal como un sitio de unión funcional para la proteína reguladora de la transcripción o la proteína reguladora de la traducción, un marco abierto de lectura aguas arriba, secuencias de unión ribosómica, un sitio de inicio de la transcripción, un sitio de inicio de la traducción y/o una secuencia de nucleótidos que codifica una secuencia líder, un codón de terminación, un sitio de parada de la traducción y una región 3' no traducida. Se incluyen en la invención promotores constitutivos o inducibles como se conocen en la técnica. Los promotores pueden ser promotores ya sea de origen natural o promotores híbridos que combinan elementos de más de un promotor. Las secuencias promotoras incluidas en la presente invención pueden ser nativas para el organismo de interés o pueden derivar de una fuente alternativa, donde la región es funcional en el organismo elegido. La elección del promotor diferirá dependiendo del hospedador previsto. Por ejemplo, los promotores que podrían usarse para la expresión en plantas incluyen promotores vegetales tales como: ejemplos de promotores vegetales constitutivos de los cuales se incluyen el promotor vegetal CaMV35S, el promotor vegetal CaMV19S, el promotor vegetal FMV34S, el promotor vegetal de badnavirus baciliforme de la caña de azúcar, el promotor vegetal CsVMV, el promotor vegetal de actina de *Arabidopsis* ACT2/ACT8, el promotor vegetal UBQ1 de ubiquitina de *Arabidopsis*, el promotor vegetal BTH6 de tionina de la hoja de cebada y el promotor vegetal de actina de arroz; ejemplos de promotores vegetales específicos de tejido de los cuales se incluyen el promotor vegetal de la proteína de almacenamiento faseolina de judía, el promotor vegetal DLEC, el promotor vegetal PHS β , el promotor vegetal de la proteína de almacenamiento zeína, el promotor vegetal de gamma conglutina de la soja, el promotor vegetal del gene AT2S1, el promotor vegetal de actina ACT11 de *Arabidopsis*, el promotor vegetal *napA* de *Brassica napus* y el promotor vegetal del gen de patatina de la patata; y

ejemplos de promotores vegetales inducibles de los cuales se incluyen un promotor vegetal inducible por la luz derivado del gen *rbcS* de guisante, un promotor vegetal del gen *rbcS* de alfalfa, los promotores vegetales DRE, MYC y MYB que son activos en la sequía; los promotores vegetales INT, INPS, *prxEa*, Ha hsp17.7G4 y RD21 activos con salinidad elevada y estrés osmótico, y los promotores vegetales hsr203J y str246C activos en el estrés patógeno.

Como alternativa, los promotores que podrían usarse para la expresión en mamíferos incluyen el promotor de metalotioneína, que puede inducirse en respuesta a metales pesados tales como cadmio, el promotor de acción β , así como promotores víricos tales como el promotor del antígeno T grande SV40, el promotor temprano inmediato (IE, por sus siglas en inglés) de citomegalovirus (CMV) humano, el promotor LTR del virus del sarcoma de Rous, el promotor de adenovirus o un promotor del VPH, en particular también puede usarse la región reguladora aguas arriba (URR, por sus siglas en inglés) del VPH. Todos estos promotores están bien descritos y están fácilmente disponibles en la técnica.

Las construcciones sintéticas de la presente invención también pueden comprender una secuencia no traducida 3'. Una secuencia no traducida 3' se refiere a aquella porción de un gen que comprende un segmento de ADN que contiene una señal de poliadenilación y cualquier otra señal reguladora capaz de efectuar el procesamiento del ARNm o la expresión génica. La señal de poliadenilación se caracteriza por efectuar la adición de segmentos de ácido poliadenílico al extremo 3' del precursor de ARNm. Las señales de poliadenilación se reconocen habitualmente por la presencia de homología con la forma canónica 5' AATAAA-3', aunque las variaciones no son infrecuentes. La secuencia de ADN reguladora no traducida 3' preferentemente incluye de aproximadamente 50 a 1000 pares de bases de nucleótidos y puede contener secuencias de terminación de la transcripción y de la traducción además de una señal de poliadenilación y cualquier otra señal reguladora capaz de efectuar el procesamiento de ARNm o la expresión génica.

En realizaciones específicas, las construcciones sintéticas contienen adicionalmente un gen marcador seleccionable para permitir la selección de un organismo o un precursor del mismo que contenga una construcción sintética. Los genes de selección son bien conocidos en la técnica y serán compatibles para la expresión en la célula u organismo de interés, o un progenitor o precursor de los mismos.

En algunas realizaciones, las construcciones sintéticas de la invención están en forma de vectores víricos, tales como virus de simio 40 (VS40) o virus del papiloma bovino (VPB), que tienen la capacidad de replicarse como elementos extracromosómicos (*Eukaryotic Viral Vectors*, Cold Spring Harbor Laboratory, Gluzman ed., 1982; Sarver et al., 1981, *Mol. Cell Biol.* 1: 486). Los vectores víricos incluyen genomas de retrovirus (lentivirus), virus adenoasociados (véanse, por ejemplo, Okada, 1996, *Gene Ther.* 3: 957-964; Muzyczka, 1994, *J. Clin. Invest.* 94: 1351; las Patentes de los EE.UU. N.º 6.156.303; 6.143.548, 5.952.221, que describen vectores de VAA; véanse también las Patentes de los EE.UU. N.º 6.004.799; 5.833.993), adenovirus (véanse, por ejemplo, las Patentes de los EE.UU. N.º 6.140.087; 6.136.594; 6.133.028; 6.120.764), reovirus, herpesvirus, rotavirus etc., modificados para introducir y dirigir la expresión de un polinucleótido o transgén en las células. Los vectores retrovíricos pueden incluir aquellos basados en el virus de la leucemia murina (véase, por ejemplo, la Patente de los EE.UU. N.º 6.132.731), el virus de la leucemia de simio gibón (véase, por ejemplo, la Patente de los EE.UU. N.º 6.033.905), el virus de la inmunodeficiencia de simio, el virus de la inmunodeficiencia humana (véase, por ejemplo, la Patente de los EE.UU. N.º 5.985.641) y combinaciones de los mismos.

Los vectores también incluyen aquellos que entregan genes eficientemente a células de animales *in vivo* (por ejemplo, células madre) (véase, por ejemplo, la Patente de los EE.UU. N.º 5.821.235 y 5.786.340; Croyle et al., 1998, *Gene Ther.* 5: 645; Croyle et al., 1998, *Pharm. Res.* 15: 1348; Croyle et al., 1998, *Hum. Gene Ther.* 9: 561; Foreman et al., 1998, *Hum. Gene Ther.* 9: 1313; Wirtz et al., 1999, *Gut* 44: 800). Los vectores víricos adenovíricos y adenoasociados adecuados para la entrega *in vivo* se describen, por ejemplo, en las Patentes de los EE.UU. N.º 5.700.470, 5.731.172 y 5.604.090. Los vectores adicionales adecuados para la entrega *in vivo* incluyen vectores del virus del herpes simple (véase, por ejemplo, la Patente de los EE.UU. N.º 5.501.979), vectores retrovíricos (véase, por ejemplo, las Patentes de los EE.UU. N.º 5.624.820, 5.693.508 y 5.674.703; y el documento WO92/05266 y el documento WO92/14829), vectores del virus del papiloma bovino (VPB) (véase, por ejemplo, la Patente de los EE.UU. N.º 5.719.054), vectores basados en el CMV (véase, por ejemplo, la Patente de los EE.UU. N.º 5.561.063) y vectores de parvovirus, rotavirus y virus de Norwalk. Los vectores lentivíricos son útiles para infectar tanto células en división como células que no están en división (véase, por ejemplo, la Patente de los EE.UU. N.º 6.013.516).

Los vectores para la expresión en células de insectos usan habitualmente variaciones recombinantes de baculovirus y otros nucleopolihedrovirus, por ejemplo, vectores de nucleopolihedrovirus de *Bombyx mori* (véase, por ejemplo, Choi, 2000, *Arch. Virol.* 145: 171-177). Por ejemplo, se usan células de lepidópteros y coleópteros para replicar baculovirus para promover la expresión de genes extraños llevados por baculovirus, por ejemplo, se infectan células de *Spodoptera frugiperda* con virus de la poliedrosis nuclear de *Autographa californica* recombinante (AcNPV) que lleva una secuencia codificante heteróloga, por ejemplo, de ser humano, (véase, por ejemplo, Lee, 2000, *J. Virol.* 74: 11873-11880; Wu, 2000, *J. Biotechnol.* 80: 75-83). Véase, por ejemplo, la Patente de los EE.UU. N.º 6.143.565, que describe el uso del polidnavirus de la avispa parásita *Glyptapanteles indiensis* para integrar de manera estable ácido nucleico en el genoma de líneas celulares de insectos lepidópteros y coleópteros. Véanse también, las Patentes de los EE.UU. N.º 6.130.074; 5.858.353; 5.004.687.

Los vectores de expresión capaces de expresar proteínas en plantas son bien conocidos en la técnica, e incluyen, por ejemplo, vectores de *Agrobacterium* spp., virus X de la patata (véase, por ejemplo, Angell, 1997, *EMBO J.* 16: 3675-3684), virus del mosaico del tabaco (véase, por ejemplo, Casper, 1996, *Gene* 173: 69-73), virus del enanismo arbustivo del tomate (véase, por ejemplo, Hillman, 1989, *Virology* 169:42-50), virus del mosaico del tabaco (véase, por ejemplo, Dolja, 1997, *Virology* 234: 243-252), virus del mosaico de la judía dorada (véase, por ejemplo, Morinaga, 1993, *Microbiol. Immunol.* 37: 471-476), virus del mosaico de la coliflor (véase, por ejemplo, Cecchini, 1997, *Mol. Plant. Microbe Interact.* 10: 1094-1101), elemento transponible Ac/Ds del maíz (véase, por ejemplo, Rubin, 1997, *Mol. Cell Biol.* 17: 6294-6302; Kunze, 1996, *Curr. Top. Microbiol. Immunol.* 204: 161-194) y elemento transponible supresor-mutador (Spm) del maíz (véase, por ejemplo, Schlappi, 1996, *Plant Mol. Biol.* 32: 717-725); y derivados de los mismos.

La invención contempla adicionalmente células u organismos que contienen en su interior las construcciones sintéticas de la invención o, como alternativa, partes, precursores, células o tejidos producidos por los métodos que se describen en el presente documento. En este sentido, se apreciará que el sistema de construcciones de la presente invención es aplicable a hospedadores procariontes, así como eucariotes e incluye, por ejemplo, organismos unicelulares y organismos multicelulares, tales como, pero no limitados a levadura, plantas y animales, incluyendo animales vertebrados tales como mamíferos, reptiles, peces, aves, etc, así como animales invertebrados tales como metazoarios, esponjas, gusanos, moluscos, nematodos, crustáceos, equinodermos etc. En ciertas realizaciones, el sistema de construcciones se usa para determinar la eficiencia de traducción de diferentes codones sinónimos en células vegetales o células animales o para determinar la preferencia fenotípica de diferentes codones sinónimos en plantas y mamíferos.

Los ejemplos ilustrativos de organismos eucariotes incluyen, pero no se limitan a, hongos tales como levaduras y hongos filamentosos, incluyendo especies de *Aspergillus*, *Trichoderma* y *Neurospora*; hospedadores animales, incluyendo animales vertebrados cuyos ejemplos ilustrativos incluyen peces (por ejemplo, salmón, trucha, tilapia, atún, carpa, platija, halibut, pez espada, bacalao y pez cebra), aves (por ejemplo, pollos, patos, codornices, faisanes y pavos y otros faisanes asiáticos o aves de caza) y mamíferos (por ejemplo, perros, gatos, caballos, vacas, búfalos, ciervos, ovejas, conejos, roedores tales como ratones, ratas, hámsteres y cobayas, cabras, cerdos, primates, mamíferos marinos, incluyendo delfines y ballenas, así como estirpes celulares, tales como estirpes celulares de seres humanos o de otro mamíferos de cualquier tipo de tejido o de células madre (por ejemplo, COS, NIH 3T3 CHO, BHK, 293 o células HeLa) y células madre, incluyendo células madre pluripotentes y no pluripotentes y embrionarias y cigotos no humanos), así como animales invertebrados cuyos ejemplos ilustrativos incluyen nematodos (cuyos géneros representativos incluyen aquellos que infectan a animales tales como, pero no limitados a, *Ancylostoma*, *Ascaridia*, *Ascaris*, *Bunostomum*, *Caenorhabditis*, *Capillaria*, *Chabertia*, *Cooperia*, *Dictyocaulus*, *Haemonchus*, *Heterakis*, *Nematodirus*, *Oesophagostomum*, *Ostertagia*, *Oxyuris*, *Parascaris*, *Strongylus*, *Toxascaris*, *Trichuris*, *Trichostrongylus*, *Tflichonema*, *Toxocara*, *Uncinaria* y aquellos que infectan plantas tales como, pero no limitados a *Bursaphelenchus*, *Criconeuriella*, *Diiylenchus*, *Ditylenchus*, *Globodera*, *Helicotylenchus*, *Heterodera*, *Longidorus*, *Meloidogyne*, *Nacobbus*, *Paratylenchus*, *Pratylenchus*, *Radopholus*, *Rotelynychus*, *Tylenchus* y *Xiphinerna*) y otros gusanos, *Drosophila* y otros insectos (tales como de las familias Apidae, Curculionidae, Scarabaeidae, Tephritidae, Tortricidae, entre otros, cuyos órdenes representativos incluyen Coleoptera, Diptera, Lepidoptera y Homoptera).

En ciertas realizaciones, el sistema de construcciones se usa para determinar la eficacia de traducción o la preferencia fenotípica de diferentes codones sinónimos en plantas o células de plantas (por ejemplo, una planta que se selecciona convenientemente entre monocotiledóneas, dicotiledóneas y gimnospermas). La planta puede ser una planta ornamental o planta de cultivo. Los ejemplos ilustrativos de plantas ornamentales incluyen, pero no se limitan a, *Malus* spp, *Crataegus* spp, *Rosa* spp., *Betula* spp, *Sorbus* spp, *Olea* spp, *Nerium* spp, *Salix* spp, *Populus* spp. Los ejemplos ilustrativos de plantas de cultivo incluyen especies de plantas que se cultivan con el fin de producir un producto cosechable tales como, pero no limitadas a, *Abelmoschus esculentus* (ocra), *Acacia* spp., *Agave fourcroydes* (henequén), *Agave sisalana* (sisal), *Albizia* spp., *Allium fistulosum* (cebollita), *Allium sativum* (ajo), *Allium* spp. (cebollas), *Alpinia galanga* (galanga mayor), *Amaranthus caudatus*, *Amaranthus* spp., *Anacardium* spp. (anacardo), *Ananas comosus* (piña), *Anethum graveolens* (eneldo), *Annona cherimola* (chirimoya), *Apios americana* (patata de la India), *Arachis hypogaea* (cacahuete), *Arctium* spp. (bardana), *Artemisia* spp. (ajenjo), *Aspalathuslinearis* (té Rooibos), *Athertonia diversifolia*, *Atriplex nummularia* (tiple), *Averrhoa carambola* (carambola), *Azadirachta indica* (nim), *Backhousia* spp., *Bambusa* spp. (bambú), *Beta vulgaris* (remolacha azucarera), *Boehmeria nivea* (ramio), *col china*, *Boronia megastigma* (boronia dulce), *Brassica carinata* (mostaza de Abisinia), *Brassica juncea* (mostaza de la India), *Brassica napus* (colza), *Brassica oleracea* (repollo, brócoli), *Brassica oleracea* var Albogabra (brócoli chino), *Brassica parachinensis* (choi sum), *Brassica pekinsis* (Bok Wong o repollo chino), *Brassica* spp., *Burcella obovata*, *Cajanus cajan* (gandul), *Camellia sinensis* (té), *Cannabis sativa* (cáñamo no fármaco), *Capsicum* spp., *Carica* spp. (papaya), *Carthamus tinctorius* (cártamo), *Carum carvi* (comino), *Cassinia* spp., *Castanospermum australe* (judía negra), *Casuarina cunninghamiana* (roble de río), *Ceratonia siliqua* (algarroba), *Chamaemelum nobile* (manzanilla), *Chamelaucium* spp. (Cera de Geraldton), *Chenopodium quinoa* (quinoa), *crisantemo* (Tanacetum), *cinerariifolium* (piretro), *Cicer arietinum* (garbanzo), *Cichorium intybus* (achicoria), *Clematis* spp., *Clianthus formosus* (guisante de desierto), *Cocos nucifera* (coco), *Coffea* spp. (café), *Colocasia esculenta* (taro), *Coriandrurn sativum* (cilantro), *Crambe abyssinica* (crambe), *Crocus sativus* (azafrán), *Cucurbita foetidissima* (calabaza de búfalo), *Cucurbita* spp. (calabaza), *Cyamopsis tetragonoloba* (guar), *Cymbopogon* spp.

(citronela), *Cytisus proliferus* (tagasaste), *Daucus carota* (zanahoria), *Desmanthus* spp., *Dioscorea esculenta* (ñame asiático), *Dioscorea* spp. (ñame), *Diospyros* spp. (caqui), *Doronicum* sp., *Echinacea* spp., *Eleocharis dulcis* (castaña de agua), *Eleusine coracana* (mijo), *Emanthus arundinaceus*, *Eragrostis tef* (tef), *Erianthus arundinaceus*, *Eriobotrya japonica* (nispero), *Eucalyptus* spp., *Eucalyptus* spp. (gil mallee), *Euclea* spp., *Eugenia malaccensis* (jumba),
5 *Euphorbia* spp., *Euphorbia Longana* (ojo de dragón), *Eutrema wasabi* (wasabi), *Fagopyrum esculentum* (trigo sarraceno), *Festuca arundinacea* (festuca alta), *Ficus* spp. (fig), *Flacourtia inermis*, *Flindersia grayliana* (arce de Queensland), *Foeniculum Olearia*, *Foeniculum vulgare* (hinojo), *Garcinia mangostana* (mangostán), *Glycine latifolia*, *Glycine max* (semilla de soja), *Glycine max* (soja verdura), *Glycyrrhiza glabra* (regaliz), *Gossypium* spp. (algodón),
10 *Grevillea* spp., *Grindelia* spp., *Guizotia abyssinica* (nug), *Harpagophyllum* sp., *Helianthus annuus* (girasol alto oleico), *Helianthus annuus* (girasol monosun), *Helianthus tuberosus* (pataca), *Hibiscus cannabinus* (kenaf), *Hordeum bulbosum*, *Hordeum* spp. (cebada c rea), *Hordeum vulgare* (cebada), *Hordeum vulgare* subsp. spontaneum, *Humulus lupulus* (l pulo), *Hydrastis canadensis* (sello de oro), *Hymenachne* spp., *Hyssopus officinalis* (hisopo), *Indigofera* spp., *Inga edulis* (guama), *Inocarpus tugiter*, *Ipomoea batatas* (boniato), *Ipomoea* sp. (kang kong), *Lablab purpureus* (lablab blanco), *Lactuca* spp. (lechuga), *Lathyrus* spp. (guisante de olor), *Lavandula* spp. (lavanda), *Lente* spp. (lentejas), *Lesquerella* spp. (bladderpod), *Leucaena* spp., *Lilium* spp., *Limnanthes* spp. (hierba de la pradera),
15 *Linum usitatissimum* (lino), *Linum usitatissimum* (lino), *Linum usitatissimum* (Linola.TM.), *Litchi chinensis* (lichi), *Lotus corniculatus* (loto de los prados), *Lotus pedunculatus*, *Lotus* sp., *Luffa* spp., *Lunaria annua* (planta de la plata), *Lupinus mutabilis* (altramuz perla), *Lupinus* spp. (altramuz), *Macadamia* spp., *Mangifera indica* (mango), *Manihot esculenta* (yuca), *Medicago* spp. (alfalfa), *Medicago* spp., *Melaleuca* spp. ( rbol de t ), *Melaleuca uncinata* (broza),
20 *Mentha tasmanica*, *Mentha spicata* (hierbabuena), *Mentha piperita* X (menta), *Momordica charantia* (mel n amargo), *Musa* spp. (pl tano), *Myrciaria cauliflora* (jaboticaba), *Myrothamnus flabellifolia*, *Nephelium lappaceum* (rambut n), *Nerine* spp., *Ocimum basilicum* (albahaca), *Oenanth  javanica* (apio de agua), *Oenothera biennis* (onagra), *Olea europaea* (olivo), *Olearia* sp., *Origanum* spp. (mejorana, or gano), *Oryza* spp. (arroz), *Oxalis tuberosa* (oca), *Ozothamnus* spp. (flor de arroz), *Pachyrrhizus ahipa* (jicama), *Panax* spp. (ginseng), *Panicum miliaceum* (mijo com n), *Papaver* spp. (amapola), *Parthenium argentatum* (guayule), *Passiflora* sp., *Paulownia tomentosa* ( rbol princesa), *Pelargonium graveolens* (geranio rosa), *Pelargonium* sp., *Pennisetum americanum* (junco o mijo perla),
25 *Persoonia* spp., *Petroselinum crispum* (perejil), *Phacelia tanacetifolia* (tanaceto), *Phalaris canariensis* (alpiste), *Phalaris* sp., *Phaseolus coccineus* (ayocote), *Phaseolus lunatus* (jud  lima), *Phaseolus* spp., *Phaseolus vulgaris* (jud  culinaria), *Phaseolus vulgaris* (jud  blanca), *Phaseolus vulgaris* (jud  roja), *Pisum sativum* (guisante),
30 *Plantago ovata* (psyllium), *Polygonum minus*, *Polygonum odoratum*, *Prunus mume* (albaricoque japones ), *Psidium guajava* (guayaba), *Psophocarpus tetragonolobus* (jud  alada), *Pyrus* spp. (nashi), *Raphanus satulus* (r bano blanco largo o Daikon), *Rhagodia* spp. (caramillo), *Ribes nigrum* (grosellero negro), *Ricinus communis* (ricino), *Rosmarinus officinalis* (romero), *Rungia klossii* (rungia), *Saccharum officinarum* (ca a de az car), *Salvia officinalis* (salvia), *Salvia sclarea* (salvia peluda), *Salvia* sp., *Sandersonia* sp., *Santalum acuminatum* (quandong dulce),
35 *Santalum* spp. (madera de s ndalo), *Sclerocarya caffra* (marula), *Scutellaria galericulata* (escutelaria), *Secale cereale* (centeno), *Sesamum indicum* (s samo), *Setaria italica* (mijo de cola de zorro), *Simmondsia* spp. (jojoba), *Solanum* spp., *Sorghum alnum* (sorgo), *Stachys betonica* (madera bret nica), *Stenanthemum scortechenii*, *Strychnos cocculoides* (naranja de los monos), *Stylosanthes* spp. (stylo), *Syzygium* spp., *Tasmanica lanceolata* (pimienta de la mont a), *Terminalia karnbachii*, *Theobroma cacao* (cacao), *Thymus vulgaris* (tomillo), *Toona australis* (cedro rojo), *Trifolium* spp. (tr boles), *Trifolium alexandrinum* (tr bol de Alejandr a), *Trifolium resupinatum* (tr bol persa), *Triticum* spp., *Triticum tauschii*, *Tylosema esculentum* (jud  Morama), *Valeriana* sp. (valeriana),
40 *Vernonia* spp., *Vetiver zizanioides* (pasto vetiver), *Vicia benghalensis* (vicia p rpura), *Vicia faba* (haba), *Vicia narbonensis* (jud  Narbon), *Vicia sativa*, *Vicia* spp., *Vigna aconitifolia* (jud  moth), *Vigna angularis* (jud  adzuki), *Vigna mungo* (gramo negro), *Vigna radiata* (jud  mungo), *Vigna* spp., *Vigna unguiculata* (caupi), *Vitis* spp. (uvas),
45 *Voandzeia subterranea* (cacahuete bambarra), *Triticosecale* (triticale), *Zea mays* (ma  dulce bicolor), *Zea mays* (ma ), *Zea mays* (ma  dulce), *Zea mays* subsp. mexicana (teosinte), *Zieria* spp., *Zingiber officinale* (jengibre), *Zizania* spp. (arroz silvestre), *Ziziphus jujuba* (azufa  com n). Los cultivos deseables para la pr ctica de la presente invenci n incluyen *Nicotiana tabacum* (tabaco) y cultivos hort colas, tales como, por ejemplo, *Ananas comosus* (pi a), *Saccharum* spp (ca a de az car), *Musa* spp (pl tano), *Lycopersicon esculentum* (tomate) y *Solanum tuberosum* (patata).
50

Las construcciones sint ticas de la presente invenci n pueden introducirse directamente *ex vivo* o en cultivo celular en una c lula de inter s o en un organismo de inter s o en una o m s de las partes de un organismo de inter s, por ejemplo, tipos celulares o tisulares (por ejemplo, un m sculo, piel, cerebro, pulm n, ri n n, p ncreas, un  rgano reproductor tal como test culos, ovarios y mama, ojos, h gado, coraz n, c lulas vasculares, ra z, hoja, flor, tallo o meristemas) o en un  rgano de un organismo de inter s. Como alternativa, las construcciones sint ticas se introducen en un progenitor de una c lula u organismo de inter s y el progenitor despu s se hace crecer o se cultivan durante un tiempo y en condiciones suficientes para diferenciarse en la c lula de inter s o producir el organismo de inter s, mediante el cual la construcci n sint tica est  contenida en la c lula de inter s o uno o m s tipos de c lulas del organismo de inter s. Las c lulas progenitoras adecuadas incluyen, pero no se limitan a, c lulas tales como c lulas madre embrionarias, c lulas inmunitarias pluripotenciales, c lulas meristem ticas y callo embrionario. En ciertas realizaciones, la construcci n sint tica se introduce en el organismo de inter s usando una v a de administraci n particular (por ejemplo, para los mam feros, por v a oral, parenteral (por ejemplo, intravenosa, intramuscular, intraperitoneal, intraventricular, intraarticular), mucosal (por ejemplo, intranasal, intrapulmonar, oral, bucal, sublingual, rectal, intravaginal), d rmica (t pica, subcut nea, transd rmica); para las plantas, la administraci n a las flores, meristemas, ra z, hojas o tallos). Los practicantes de la t cnica reconocer n que la v a de administraci n
55
60
65

será diferente dependiendo de la elección del organismo de interés y el fenotipo buscado. En algunas realizaciones relativas a la determinación de la preferencia fenotípica, las construcciones sintéticas se introducen convenientemente en el mismo o el correspondiente sitio del organismo o parte del mismo. En otras realizaciones, las construcciones sintéticas se introducen en una célula del organismo de interés (por ejemplo, células autólogas) o en una célula que es compatible con el organismo de interés (por ejemplo, células singénicas o alogénicas) y la célula modificada genéticamente producida de este modo se introduce en el organismo de interés en un sitio seleccionado o en una parte de ese organismo.

Las construcciones sintéticas de la presente invención pueden introducirse en una célula u organismo de interés o parte de los mismos usando cualquier método adecuado y el tipo de método empleado será diferente dependiendo del tipo celular, parte y/u organismo de interés previstos. Por ejemplo, se han descrito cuatro clases generales de los métodos para la entrega de moléculas de ácido nucleico en células: (1) métodos químicos, tales como precipitación con fosfato de calcio, precipitación mediada por polietilenglicol (PEG) y lipofección; (2) métodos físicos tales como microinyección, electroporación, métodos de aceleración e infiltración al vacío; (3) métodos basados en vectores tales como transformación bacteriana y vírica mediada por vectores; y (4) métodos mediados por receptores. Las técnicas de transformación que pertenecen a estas y otras clases son bien conocidas por los trabajadores en la técnica y continuamente se están dando a conocer nuevas técnicas. La elección particular de una tecnología de transformación estará determinada por su eficiencia para transformar ciertas especies hospedadoras, así como por la experiencia y la preferencia de la persona que pone en práctica la invención con una metodología particular de elección. Será evidente para el experto que la elección particular de un sistema de transformación para introducir una construcción sintética de la invención en las células no es esencial ni una limitación para la invención, a condición de que alcance un nivel aceptable de transferencia de ácido nucleico. Por tanto, las construcciones sintéticas se introducen en los tejidos o células hospedadoras mediante cualquier número de vías, incluyendo la infección vírica, la infección por fagos, la microinyección, la electroporación o la fusión de vesículas, la lipofección, la infección por *Agrobacterium tumefaciens* o *A. rhizogenes* o la fusión de protoplastos. También puede usarse la inyección a chorro para la administración intramuscular (como se describe por ejemplo por Furth et al., 1992, *Anal Biochem* 205: 365-368). Las construcciones sintéticas pueden aplicarse como recubrimiento sobre microproyectiles y se entregan en una célula hospedadora o en el tejido mediante un dispositivo de bombardeo de partículas o "pistola de genes" (véase, por ejemplo, Tang et al., 1992, *Nature* 356: 152-154). Como alternativa, las construcciones sintéticas pueden alimentarse directamente o inyectarse en, un organismo hospedador o pueden introducirse en una célula (es decir, intracelularmente) o pueden introducirse extracelularmente en una cavidad, espacio intersticial, en la circulación de un organismo, introducirse por vía oral, etc. Los métodos para la introducción oral incluyen la mezcla directa de las construcciones sintéticas con el alimento del organismo. En ciertas realizaciones, se emplea un protocolo de administración hidrodinámica de ácido nucleico (por ejemplo, véase Chang et al., 2001, *J. Virol.* 75: 3469-3473; Liu et al., 1999, *Gene Ther.* 6: 1258-1266; Wolff et al., 1990, *Science* 247: 1465-1468; Zhang et al., 1999, *Hum. Gene Ther.* 10: 1735-1737; y Zhang et al., 1999, *Gene Ther.* 7: 1344-1349). Otros métodos de entrega de ácido nucleico incluyen, pero no se limitan a, la transferencia mediada por liposomas, la entrega de ADN desnudo (inyección directa) y la transferencia mediada por receptor (complejo ligando-ADN).

4. Métodos de determinación de la eficiencia de traducción o la preferencia fenotípica de codones sinónimos

El sistema de construcciones de la presente invención puede usarse para comparar la eficiencia de traducción de diferentes codones sinónimos en las células de un tipo particular o para comparar la eficiencia de traducción de codones sinónimos individuales entre diferentes tipos celulares. Sin desear quedar ligado por ninguna teoría o modo de operación particular, se cree que los niveles de proteína indicadora producida en una célula de interés a partir de construcciones sintéticas individuales son sensibles a la abundancia intracelular de las especies de iso-ARNt que corresponden a el o los codones de interrogación en las secuencias de codificación correspondientes y, por tanto, proporcionan una correlación directa de la preferencia por o de la eficiencia en la traducción de un codón dado. Esto significa, por ejemplo, que si el nivel de la proteína indicadora obtenida en una célula del mismo tipo que una célula de interés, a la que se le proporciona una construcción sintética que tiene al menos un primer codón de interrogación, es mayor que el nivel producido en una célula del mismo tipo que la célula de interés, a la que se le proporciona otra construcción sintética que tiene al menos un segundo codón de interrogación (es decir, en la que el primer codón o codones de interrogación son diferentes de, pero sinónimos de, el segundo codón o codones de interrogación), entonces puede deducirse que el primer codón de interrogación tiene una mayor eficiencia de traducción que el segundo codón de interrogación en la célula de interés. Los métodos para medir los niveles de proteína indicadora son bien conocidos en la técnica e incluyen, pero no se limitan a, inmunoensayos tales como los ensayos de transferencia Western, ELISA y RIA, ensayos de proteínas quimioluminiscentes tales como los ensayos de luciferasa, ensayos enzimáticos, tales como ensayos que miden la actividad β -galactosidasa o cloranfenicol acetil transferasa (CAT), así como ensayos fluorométricos que miden la fluorescencia asociada a una proteína fluorescente. En algunas realizaciones, las diferentes construcciones sintéticas se introducen por separado en células diferentes. En otras realizaciones, las diferentes construcciones sintéticas se introducen en la misma célula (por ejemplo, cuando los polinucleótidos indicadores comprenden secuencias codificantes auxiliares que codifican un marcador, como se describe en el presente documento).

Con respecto a la expresión diferencial del polinucleótido indicador entre diferentes tipos celulares, se apreciará que si el nivel de la proteína indicadora obtenida en un primer tipo celular al que se le proporciona una construcción

sintética que tiene al menos un codón de interrogación es mayor que el nivel obtenido en un segundo tipo celular al que se le proporciona la misma construcción sintética, entonces puede deducirse que el codón de interrogación tiene una mayor eficiencia de traducción en el primer tipo celular que en el segundo tipo celular.

Las eficiencias de traducción de diferentes codones sinónimos determinados de este modo entonces se comparan normalmente para proporcionar un orden clasificado de codones sinónimos individuales de acuerdo con su preferencia para la traducción en la célula o células de interés. Un experto habitual en la materia de este modo será capaz de determinar una "tabla de las eficiencias de traducción de los codones" para cada aminoácido. Entonces, la comparación de los codones sinónimos dentro de una tabla eficiencia de traducción codón puede usarse para identificar codones para adaptar un polinucleótido sintético para modular el nivel de un polipéptido codificado que se expresa en un tipo celular de interés o para expresar diferencialmente un polipéptido codificado entre diferentes tipos celulares.

En otras realizaciones, el sistema de construcciones se usa para comparar la preferencia de diferentes codones sinónimos para producir un fenotipo seleccionado en un organismo de interés o parte del mismo (es decir, la "preferencia fenotípica"). En estas realizaciones, las construcciones sintéticas se usan para determinar la influencia del codón o codones de interrogación en el fenotipo o clase de fenotipo presentado por el organismo o parte del mismo en respuesta a la proteína asociada a fenotipo producida por las construcciones sintéticas. Esto significa, por ejemplo, que si la cualidad del fenotipo presentado por el organismo o parte del mismo al que se le proporciona una construcción sintética que tiene al menos un primer codón de interrogación es mayor que la cualidad del fenotipo presentado por el organismo o parte del mismo al que se le proporciona una construcción sintética que tiene al menos un segundo codón de interrogación (es decir, en el que el primer codón de interrogación es diferente, pero sinónimo, del segundo codón de interrogación), entonces puede deducirse que el organismo de interés o parte del mismo tiene una mayor preferencia por el primer codón de interrogación que por el segundo codón de interrogación con respecto a la cualidad del fenotipo producido. Dicho de otro modo, el primer codón de interrogación tiene una preferencia fenotípica mayor que el segundo codón de interrogación en el organismo de interés o parte del mismo.

De acuerdo con la presente invención, las construcciones sintéticas individuales se introducen en organismos de ensayo que se seleccionan preferentemente entre organismos de la misma especie que el organismo u organismos de interés que se relacionan con el organismo de interés o en partes de ensayo de dichos organismos. Los organismos relacionados son en general especies dentro del mismo phylum, preferentemente especies dentro del mismo subphylum, más preferentemente especies dentro de superclase, incluso más preferentemente especies dentro de la misma clase, incluso más preferentemente especies dentro del mismo orden e incluso aún más preferentemente especies dentro del mismo género. Por ejemplo, si el organismo de interés es un ser humano, una especie relacionada se selecciona convenientemente entre ratón, vaca, perro o gato, que pertenecen a la misma clase que el ser humano, o un chimpancé, que pertenece al mismo orden que el ser humano. Como alternativa, si el organismo de interés es el plátano, el organismo relacionado puede seleccionarse entre taro, jengibre, cebolla, ajo, piña, bromelias, palmeras, orquídeas, lirios, iris y similares, que son todas plantas monocotiledóneas no gramíneas y que constituyen parientes hortícolas o botánicos.

Después de la introducción de las construcciones sintéticas dentro de los organismos de ensayo o partes de los mismos, las cualidades de sus fenotipos se determinan mediante un ensayo adecuado y después se comparan para determinar las preferencias fenotípicas relativas de los codones sinónimos. La cualidad es convenientemente una medida de la potencia, intensidad o grado del fenotipo, o la potencia, intensidad o grado relativos de dos o más rasgos fenotípicos deseados. Los ensayos para diversos fenotipos conferidos por la producción de una proteína indicadora elegida son conocidos por los expertos en la materia. Por ejemplo, la inmunidad puede someterse a ensayo por cualquier método adecuado que detecte un aumento de la capacidad de un animal para responder a antígenos extraños o específicos de la enfermedad (por ejemplo, antígenos de cáncer) es decir, aquellas células cebadas para atacar dichos antígenos se incrementan en número, actividad y capacidad para detectar y destruir esos antígenos. La potencia de la respuesta inmunitaria se mide mediante ensayos convencionales, incluyendo: la medición directa de los linfocitos de sangre periférica por medios conocidos en la técnica; ensayos de citotoxicidad de linfocitos citolíticos naturales (véase, por ejemplo, Provinciali et al (1992, *J. Immunol. Meth.* 155: 19-24), ensayos de proliferación celular (véase, por ejemplo, Vollenweider y Groseurth (1992, *J. Immunol. Meth.* 149: 133-135), inmunoensayos de células y subconjuntos inmunitarios (véase, por ejemplo, Loeffler et al. (1992, *Cytom.* 13: 169-174); Rivoltini et al. (1992, *Can. Immunol Immunother.* 34: 241-251); o ensayos cutáneos para la inmunidad mediada por células (véase, por ejemplo, Chang et al (1993, *Cancer Res.* 53: 1043-1050.) La respuesta inmunitaria potenciada también se indica por manifestaciones físicas tales como la fiebre y la inflamación, así como la curación de infecciones sistémicas y locales, y la reducción de síntomas en la enfermedad, es decir, la disminución del tamaño tumoral, el alivio de los síntomas de una enfermedad o afección, incluyendo, pero no restringida a, lepra, tuberculosis, malaria, úlceras aftosas, verrugas herpéticas y papilomatosas, gingivitis, aterosclerosis, los acompañantes del SIDA tales como el sarcoma de Kaposi, infecciones bronquiales y similares. Dichas manifestaciones físicas también pueden usarse para detectar o definir la cualidad, el fenotipo o clase de fenotipo presentado por un organismo. Como alternativa, la tolerancia a los herbicidas puede someterse a ensayo mediante el tratamiento de organismos de ensayo (por ejemplo, plantas tales como plantas de algodón), que expresan un gen de tolerancia a herbicidas (por ejemplo, gen de la proteína de tolerancia al glifosato tal como la sintasa EPSP resistente a glifosato), con un herbicida (por ejemplo, glifosato) y la determinación de la eficacia de la tolerancia a

herbicidas presentada por las plantas. Por ejemplo, cuando se determina la eficacia de construcciones sintéticas para conferir tolerancia al herbicida en algodón, la cantidad de retención de la cápsula es una medida de la eficacia y es un rasgo deseable.

Después se comparan las cualidades del fenotipo seleccionado presentado por los organismos de ensayo o por las partes de ensayo para proporcionar un orden clasificado de los codones sinónimos individuales de acuerdo con su preferencia de uso por el organismo o parte del mismo para conferir el fenotipo seleccionado. Un experto habitual en la materia ser capaz de determinar de este modo una "tabla de preferencia de codón" para cada aminoácido en el polipéptido cuya expresión transmite el fenotipo seleccionado al organismo de interés. La comparación de los codones sinónimos dentro de una tabla de preferencia de codón después puede usarse para identificar codones para adaptar un polinucleótido sintético para modular la cualidad de un fenotipo seleccionado.

5. Modificación por codones de polinucleótidos

El sistema de construcciones de la presente invención puede usarse, por tanto, para proporcionar una comparación de las eficiencias de traducción para codones sinónimos en una célula de interés o una comparación de las preferencias fenotípicas para codones sinónimos en un organismo de interés o en un organismo relacionado, o en partes del mismo. Estas comparaciones pueden usarse como base para la construcción de un polinucleótido sintético o 'modificado por codones' que difiere de un polinucleótido parental o de referencia por la sustitución de al menos un codón 'reemplazable' (también denominado en el presente documento "un primer codón") en el polinucleótido parental por un codón sinónimo que tiene una eficiencia de traducción diferente o una preferencia fenotípica diferente que las del codón reemplazable.

5.1 Modificaciones basadas en codones sinónimos con diferentes eficiencias de traducción

En algunas realizaciones, el polinucleótido sintético se construye de manera que produzca un polipéptido codificado en una célula de interés en un nivel diferente que el producido a partir de un polinucleótido parental. El método comprende seleccionar un codón reemplazable del polinucleótido parental para el reemplazo por un codón sinónimo, en el que el codón sinónimo se selecciona basándose en éste presenta una eficiencia de traducción diferente de la del codón reemplazable en una comparación de las eficiencias de traducción en la célula de interés, como se determina, por ejemplo, en la Sección 4. El codón reemplazable después se reemplaza por el codón sinónimo para construir el polinucleótido sintético.

Por tanto, pueden seleccionarse codones sinónimos para aumentar o disminuir el nivel de polipéptido que se produce en una célula de interés. Por ejemplo, cuando se desea aumentar el nivel de polipéptido que se produce en la célula, en general es deseable usar un codón sinónimo cuya eficiencia de traducción sea al menos aproximadamente un 10 %, un 15 %, un 20 %, un 25 %, un 30 %, un 35 %, un 40 %, un 45 %, un 50 %, un 55 %, un 60 %, un 65 %, un 70 %, un 75 %, un 80 %, un 85 %, un 90 % o un 95 % superior o al menos aproximadamente 2, 3, 4, 5, 6, 7, 8, 9, 10, 50 o 100 veces superior a la eficiencia de traducción del codón reemplazable. Como alternativa, cuando se desea disminuir el nivel de polipéptido que se produce en la célula, en general es deseable usar un codón sinónimo cuya eficiencia de traducción no sea superior al 95 %, al 90 %, al 85 %, al 80 %, al 75 %, al 70 %, al 65 %, al 60 %, al 55 %, al 50 %, al 45 %, al 40 %, al 35 %, al 30 %, al 25 %, al 20 %, al 15 %, al 10 %, al 5 %, al 1 %, al 0,5 %, al 0,1 %, al 0,05 % o al 0,01 % de la eficiencia de traducción del codón reemplazable.

En general, la diferencia en el nivel de polipéptido producido en la célula de un polinucleótido sintético en comparación con el producida a partir de un polinucleótido parental depende del número de codones reemplazables que son reemplazados por codones sinónimos y de la diferencia en las eficiencias de traducción entre los codones reemplazables y los codones sinónimos en la célula de interés. Dicho de otro modo, a menor número de dichos reemplazos, y/o cuanto menor sea la diferencia en las eficiencias de traducción entre los codones sinónimos y los reemplazables, menor será la diferencia en la producción de proteínas entre el polinucleótido sintético y el polinucleótido parental. Por el contrario, a mayor número de dichas sustituciones, y/o cuanto mayor sea la diferencia en las eficiencias de traducción entre los codones sinónimos y los reemplazables, mayor será la diferencia en la producción de proteínas entre el polinucleótido sintético y el polinucleótido parental.

En consecuencia, cuando se desea aumentar o disminuir el nivel de polipéptido producido en la célula de interés, en general es deseable pero no necesario reemplazar todos los codones reemplazables del polinucleótido parental con codones sinónimos que tengan eficiencias de traducción mayores o menores en la célula de interés, según sea el caso, que los codones reemplazables. Pueden lograrse cambios en la expresión incluso con una sustitución parcial. Normalmente, la etapa de sustitución afecta al menos aproximadamente al 5 %, al 10 %, al 15 %, al 20 %, al 25 %, al 30 %, al 35 %, al 40 %, al 50 %, al 60 %, al 70 %, al 80 %, al 90 %, al 95 %, al 99 % o más de los codones reemplazables del polinucleótido parental. Convenientemente, el número de, y la diferencia en la eficiencia de traducción entre, los codones reemplazables y los codones sinónimos se seleccionan de manera que el polipéptido seleccionado se produzca a partir del polinucleótido sintético en la célula en un nivel que sea al menos aproximadamente un 10 %, un 15 %, un 20 %, un 25 %, un 30 %, un 35 %, un 40 %, un 45 %, un 50 %, un 55 %, un 60 %, un 65 %, un 70 %, un 75 %, un 80 %, un 85 %, un 90 % o un 95 % superior, o incluso al menos aproximadamente 2, 3, 4, 5, 6, 7, 8, 9, 10, 50 o 100 veces superior, o no superior al 95 %, al 90 %, al 85 %, al 80 %, al 75 %, al 70 %, al 65 %, al 60 %, al 55 %, al 50 %, al 45 %, al 40 %, al 35 %, al 30 %, al 25 %, al 20 %, al 15 %, al 10 %, al 5 %, al 1 %, al 0,5 %, al 0,1 %, al 0,05 % o al 0,01 % de la eficiencia de traducción del codón reemplazable.

al 75 %, al 70 %, al 65 %, al 60 %, al 55 %, al 50 %, al 45 %, al 40 %, al 35 %, al 30 %, al 25 %, al 20 %, al 15 %, al 10 %, al 5 %, al 1 %, al 0,5 %, al 0,1 %, al 0,05 % o al 0,01 % del nivel al que se produce el polipéptido del polinucleótido parental en la célula. En el caso de dos o más codones sinónimos que tengan eficiencias de traducción similares, se apreciará que puede usarse cualquiera de estos codones para sustituir el codón reemplazable. Generalmente, si un polinucleótido parental tiene una elección de codones de eficiencia de traducción baja y media, es preferible, en primera instancia, reemplazar algunos o más preferentemente todos, los codones de eficiencia de traducción baja por codones sinónimos que tengan eficiencias de traducción intermedias, o preferentemente altas, cuando se requiere una mayor producción de polipéptido. Normalmente, el reemplazo de codones de eficiencia de traducción bajas por intermedias o altas da como resultado un aumento sustancial en el nivel de polipéptido producido por el polinucleótido sintético construido de este modo. Sin embargo, también es preferible reemplazar algunos, o preferentemente todos, los codones de eficiencias de traducción intermedias por codones de eficiencias de traducción altas para conferir una producción óptima del polipéptido codificado.

5.2 Modificaciones basadas en codones sinónimos con preferencias fenotípicas diferentes

En otras realizaciones, el polinucleótido sintético se construye de manera que su expresión en el organismo o parte del mismo confiera un fenotipo seleccionado en ese organismo o parte del mismo, pero en una cualidad diferente que el conferido por un polinucleótido parental que codifica el mismo polipéptido. El método comprende la selección de un codón reemplazable del polinucleótido parental para el reemplazo por un codón sinónimo, en el que el codón sinónimo se selecciona basándose en que presenta una preferencia fenotípica diferente a la del primer codón en una comparación de preferencias fenotípicas en el organismo de interés o en un organismo relacionado o en una parte del mismo, como se determina en la Sección 4. El codón reemplazable después se reemplaza por el codón sinónimo para construir el polinucleótido sintético.

Por tanto, un polinucleótido parental puede modificarse con codones sinónimos de manera que la cualidad del fenotipo seleccionado conferida por el polinucleótido modificado de este modo (polinucleótido sintético) sea mayor que a partir del polinucleótido parental. En general, la diferencia entre las respectivas cualidades fenotípicas conferidas por un polinucleótido sintético y por un polinucleótido parental depende del número de primeros codones que se reemplazan por codones sinónimos y de la diferencia en la preferencia fenotípica entre los primeros codones y los codones sinónimos en el organismo de interés o parte del mismo. Dicho de otro modo, a menor número de dichos reemplazos, y/o cuanto menor sea la diferencia en la preferencia fenotípica entre los codones sinónimos y los primeros, menor será la diferencia en la cualidad fenotípica entre los polinucleótidos sintéticos y los parentales. Por el contrario, a mayor número de dichos reemplazos y/o cuanto mayor sea la diferencia en la preferencia fenotípica entre los codones sinónimos y los primeros, mayor será la diferencia en la cualidad fenotípica entre los polinucleótidos sintéticos y los parentales.

En algunas realizaciones en las que se necesita que una mayor cualidad de un fenotipo seleccionado se presentada por un organismo de interés o parte del mismo, un codón reemplazable del polinucleótido parental se selecciona convenientemente para el reemplazo por un codón sinónimo, en el que el codón sinónimo se selecciona basándose en que presenta una preferencia fenotípica mayor que el codón reemplazable en una comparación de preferencias fenotípicas en el organismo de interés o en un organismo relacionado o en una parte del mismo. En general, una preferencia fenotípica mayor se correlacionará con una mayor cualidad del fenotipo seleccionado. De este modo, en un ejemplo no limitante de dicha correlación, se considera que un codón sinónimo tiene al menos aproximadamente una preferencia fenotípica un 10 % superior que un codón reemplazable cuando la cualidad del fenotipo presentado por un organismo o parte del mismo al que se le ha proporcionado una construcción sintética que comprende el codón sinónimo como el codón de interrogación es al menos aproximadamente un 10 % superior a la cualidad del fenotipo presentado por un organismo o parte del mismo al que se le ha proporcionado una construcción sintética que comprende el codón reemplazable como el codón de interrogación. Cuando se desea aumentar la cualidad de un fenotipo, en general es deseable usar un codón sinónimo cuya preferencia fenotípica (es decir, la preferencia para conferir ese fenotipo en el organismo o en parte del mismo) sea al menos aproximadamente un 10 %, un 15 %, un 20 %, un 25 %, un 30 %, un 35 %, un 40 %, un 45 %, un 50 %, un 55 %, un 60 %, un 65 %, un 70 %, un 75 %, un 80 %, un 85 %, un 90 % o un 95 % superior o al menos aproximadamente 2, 3, 4, 5, 6, 7, 8, 9, 10, 50 o 100 veces superior que la preferencia fenotípica del codón reemplazable. En el caso de dos o más codones sinónimos que tengan preferencias fenotípicas similares, se apreciará que puede usarse cualquiera de estos codones para reemplazar el primer codón. En general, si un polinucleótido parental tiene una selección de codones de preferencias fenotípicas bajas y medias, es preferible, en primera instancia, reemplazar algunos o más preferentemente todos los codones de preferencia fenotípica baja por codones sinónimos que tengan preferencias fenotípicas intermedias o preferentemente altas. Normalmente, el reemplazo de codones de preferencia fenotípica baja por codones de preferencia fenotípica intermedia o alta da como resultado un aumento sustancial en la cualidad del fenotipo conferida por el polinucleótido sintético construido de este modo. Sin embargo, también es preferible reemplazar algunos o preferentemente todos los codones de preferencia fenotípica intermedia por codones de eficiencia de la traducción alta para conferir una cualidad óptima en el fenotipo seleccionado.

En algunas realizaciones en las que se necesita que una menor cualidad de un fenotipo seleccionado sea presentada por un organismo de interés o parte del mismo, se selecciona un codón reemplazable del polinucleótido parental para el reemplazo por un codón sinónimo, en el que el codón sinónimo se selecciona basándose en que

presenta una preferencia fenotípica más baja que el codón reemplazable en una comparación de preferencias fenotípicas en el organismo de interés o en un organismo relacionado o en una parte del mismo, como se determina, por ejemplo, de acuerdo con el método descrito en la Sección 4. Una preferencia fenotípica inferior normalmente se correlaciona con una cualidad inferior del fenotipo seleccionado. En consecuencia, en un ejemplo no limitante de dicha correlación, se considera que un codón sinónimo tiene una preferencia fenotípica al menos aproximadamente un 10 % inferior a un primer codón cuando la cualidad del fenotipo presentado por un organismo o parte del mismo al que se le ha proporcionado una construcción sintética que comprende el codón sinónimo como el codón de interrogación es al menos aproximadamente un 10 % inferior a la cualidad del fenotipo presentado por un organismo o parte del mismo a la que se ha proporcionado una construcción sintética que comprende el codón reemplazable como el codón de interrogación. Cuando se selecciona el codón sinónimo para esta realización, se prefiere que tenga una preferencia fenotípica en el organismo de interés que no sea superior a aproximadamente el 95 %, el 90 %, el 85 %, el 80 %, el 75 %, el 70 %, el 65 %, el 60 %, el 55 %, el 50 %, el 45 %, el 40 %, el 35 %, el 30 %, el 25 %, el 20 %, el 15 %, el 10 %, el 5 %, el 1 %, el 0,5 %, el 0,1 %, el 0,05 % o el 0,01 % de la preferencia fenotípica del codón reemplazable.

Es preferible pero no necesario reemplazar todos los codones reemplazables del polinucleótido parental con codones sinónimos que tengan una preferencia fenotípica superior o inferior en el organismo de interés o parte del mismo a la de los primeros codones. Por ejemplo, puede lograrse una cualidad fenotípica superior o inferior incluso con un reemplazo parcial. Normalmente, la etapa de reemplazo afecta al 5 %, al 10 %, al 15 %, al 20 %, al 25 %, al 30 %, al 35 %, al 40 %, al 50 %, al 60 %, al 70 %, al 80 %, al 90 %, al 95 %, al 99 % o más de los codones reemplazables del polinucleótido parental. En algunas realizaciones que necesitan una cualidad fenotípica superior, el número y la diferencia en la preferencia fenotípica entre los codones reemplazables y los codones sinónimos se seleccionan de manera que el polipéptido asociado a fenotipo se produzca a partir del polinucleótido sintético para conferir un fenotipo en un organismo elegido o parte de organismo en una cualidad que sea al menos aproximadamente un 10 %, un 15 %, un 20 %, un 25 %, un 30 %, un 35 %, un 40 %, un 45 %, un 50 %, un 55 %, un 60 %, un 65 %, un 70 %, un 75 %, un 80 %, un 85 %, un 90 % o un 95 % superior o incluso al menos aproximadamente 2, 3, 4, 5, 6, 7, 8, 9, 10, 50 o 100 veces superior a la cualidad del fenotipo conferida por el polinucleótido parental en el organismo o parte del mismo. Por el contrario, en algunas realizaciones que necesitan una cualidad fenotípica inferior, el número y la diferencia en la preferencia fenotípica entre los codones reemplazables y los codones sinónimos se seleccionan de manera que el polipéptido asociado a fenotipo se produzca a partir del polinucleótido sintético para conferir un fenotipo en un organismo elegido o parte del mismo en una cualidad que es no sea superior a aproximadamente el 95 %, el 90 %, el 85 %, el 80 %, el 75 %, el 70 %, el 65 %, el 60 %, el 55 %, el 50 %, el 45 %, el 40 %, el 35 %, el 30 %, el 25 %, el 20 %, el 15 %, el 10 %, el 5 %, el 1 %, el 0,5 %, el 0,1 %, el 0,05 % o el 0,01 % de la cualidad del fenotipo conferida por el polinucleótido parental en el organismo o parte del mismo.

5.3 Construcción de polinucleótidos sintéticos

El reemplazo de un codón por otro puede conseguirse usando métodos convencionales conocidos en la técnica. Por ejemplo, la modificación por codones de un polinucleótido parental puede efectuarse usando varias técnicas de mutagénesis conocidas incluyendo, por ejemplo, la mutagénesis dirigida a oligonucleótidos, la mutagénesis con oligonucleótidos degenerados y la mutagénesis específica de región. Se describen técnicas de mutagénesis *in vitro*, por ejemplo, en las Patentes de los EE.UU. N.º 4.184.917, 4.321.365 y 4.351.901 o en las secciones pertinentes de Ausubel, et al. (*CURRENT PROTOCOLS IN MOLECULAR BIOLOGY*, John Wiley & Sons, Inc., 1997) y de Sambrook, et al., (*MOLECULAR CLONING. A LABORATORY MANUAL*, Cold Spring Harbor Press, 1989). En lugar de mediante mutagénesis *in vitro*, el polinucleótido sintético puede sintetizarse *de novo* usando maquinaria fácilmente disponible como se describe, por ejemplo, en la Patente de los EE.UU. N.º 4.293.652. Sin embargo, hay que señalar que la presente invención no depende de y no se refiere a, ninguna técnica particular para construir el polinucleótido sintético.

El polinucleótido parental es convenientemente un gen natural. Sin embargo, es posible que el polinucleótido parental no sea de origen natural pero que se ha diseñado usando técnicas recombinantes. Pueden obtenerse polinucleótidos parentales a partir de cualquier fuente adecuada, tal como a partir de organismos eucariotas o procariotas, incluyendo, pero no limitados a mamíferos u otros animales y organismos patógenos tales como levaduras, bacterias, protozoos y virus.

6. Clasificación de preferencia de respuesta inmunitaria de codones en mamíferos

El sistema de construcciones de la presente invención se ha usado para determinar experimentalmente una clasificación de codones sinónimos individuales de acuerdo con su preferencia para la producción de una respuesta inmunitaria, incluyendo una respuesta inmunitaria humoral, a un antígeno en un mamífero. En consecuencia, la presente divulgación proporciona por primera vez una clasificación de preferencias de respuesta inmunitaria de codones sinónimos individuales en mamíferos. Esta clasificación se determinó usando un sistema de construcciones que comprende una serie de construcciones indicadoras que comprenden, cada una, una secuencia codificante diferente para un polipéptido antigénico (por ejemplo, un polipéptido E7 del virus del papiloma), en el que la secuencia codificante de construcciones individuales se distingue de una secuencia codificante parental (por

- ejemplo, de tipo silvestre) por que codifica el polipéptido antigénico mediante la sustitución de una sola especie de codón isoaceptor por otras especies de codón isoaceptor que están presentes en la secuencia codificante parental. En consecuencia, las secuencias codificantes de construcciones sintéticas individuales usan el mismo codón iso-ceptor "de interrogación" para codificar al menos 1, en general al menos 2, por lo general al menos 3 casos, normalmente al menos la mayoría de los casos y preferentemente cada caso de un resto de aminoácido particular en el polipéptido antigénico y construcciones sintéticas individuales difieren en la especie de codón de interrogación iso-aceptor utilizado para codificar un resto aminoácido particular en una o más posiciones diferentes en la secuencia del polipéptido. Por ejemplo, en un polipéptido antigénico que contiene varios restos de alanina, la secuencia codificante de una construcción sintética en el sistema de construcciones de la presente invención puede comprender Ala^{GCT} como codón de interrogación para cada resto de alanina codificado, mientras que la secuencia codificante de otra construcción puede comprender Ala^{GCC} como codón de interrogación para cada resto de alanina codificado y así sucesivamente. Un sistema de construcciones sintéticas ilustrativo se describe en el Ejemplo 1, que ocubre todo el conjunto de codones sinónimos que codifican aminoácidos.
- Con el fin de determinar la preferencia de respuesta inmunitaria de diferentes codones, se inmunizan mamíferos de ensayo (por ejemplo, ratones) con el sistema de construcciones sintéticas en el que los mamíferos individuales se inmunizaron con una construcción sintética diferente y la respuesta inmunitaria del hospedador (por ejemplo, la respuesta inmunitaria humoral o una respuesta inmunitaria celular) con el polipéptido antigénico se determina para cada construcción. De acuerdo con la presente divulgación, la potencia de la respuesta inmunitaria obtenida a partir de construcciones sintéticas individuales proporciona una correlación directa con la preferencia inmunitaria de un codón de interrogación correspondiente en un mamífero de ensayo. En consecuencia, cuanto más potente sea la respuesta inmunitaria producida a partir de una construcción dada en un mamífero de ensayo, mayor será la preferencia inmunitaria será del codón de interrogación correspondiente.
- En un ejemplo ilustrativo, la comparación de las preferencias de respuesta inmunitaria determinadas de acuerdo con el Ejemplo 1 con las eficiencias de traducción derivadas de valores de frecuencia de uso de codones para células de mamíferos en general como se determina por Seed (véanse las Patentes de los EE.UU. N.º de serie 5.786.464 y 5.795.737) revela varias diferencias en la clasificación de codones. Por comodidad, estas diferencias se destacan en la TABLA 9, en la que los codones 'preferidos' de Seed se destacan con un fondo de color azul, los codones 'menos preferidos' de Seed se destacan con un fondo de color verde y los codones 'no preferidos' de Seed se resaltan con un color gris fondo.

TABLA 9

aa	Uso de codones preferentes previsto por Seed para células de mamífero en general	Preferencias respuesta inmunitaria de codón determinadas experimentalmente en mamíferos de ensayo
Ala	GCC >> (GCG, GCT, GCA)	GCT > GCC > (GCA GCG)
Arg	CGC >> (CGA, CGT, AGA, AGG, CGG)	(CGA, CGC, CGT, AGA) > (AGG, CGG)
Asn	AAC >> AAT	AAC > AAT
Asp	GAC >> GAT	GAC > GAT
Cys	TGC >> TGT	TGC > TGT
Glu	(GAA, GAG)	GAA > GAG
Gln	CAG >> CAA	CAA = CAG
Gly	GGC > GGG > (GGT, GGA)	GGA > (GGG, GGT, GGC)
His	CAC >> CAT	CAC = CAT
Ile	ATC > ATT > ATA	ATC >> ATT > ATA
Leu	CTG > CTC > (TTA,CTA, CTT, TTG)	(CTG, CTC) > (CTA, CTT) >> TTG > TTA
Lys	AAG >> AAA	AAG = AAA
Phe	TTC >> TTT	TTT > TTC
Pro	CCC >> (CCG, CCA, CCT)	CCC > CCT >> (CCA, CCG)
Ser	AGC > TCC > (TCG, AGT, TCA, TCT)	TCG >> (TCT, TCA, TCC) >> (AGC, AGT)
Thr	ACC >> (ACG, ACA, ACT)	ACG > ACC >> ACA > ACT
Tyr	TAC >> TAT	TAC > TAT
Val	GTG > GTC > (GTA, GTT)	(GTG, GTC) > GTT > GTA

Como será evidentes a partir de la tabla anterior:

(i) varios codones que Seed consideró que tenían una clasificación de uso de codones superior en células de mamífero con respecto a al menos otro codón sinónimo tienen en realidad una clasificación de preferencia de respuesta inmunitaria inferior al codón sinónimo o a cada uno de los codones sinónimo (por ejemplo, Ala^{GCC} tiene una clasificación de uso de codones superior pero una clasificación de preferencia de respuesta inmunitaria inferior con respecto a Ala^{GCT}; Gly^{GGC} tiene una clasificación de uso de codones superior pero una clasificación de preferencia de respuesta inmunitaria inferior con respecto a Gly^{GGA}; Phe^{TTC} tiene una clasificación de uso de codones superior pero una clasificación de preferencia de respuesta inmunitaria inferior con respecto a Phe^{TTT}; Ser^{AGC} tiene una clasificación de uso de codones superior pero una clasificación de preferencia de respuesta inmunitaria inferior con respecto a uno cualquiera de Ser^{TCG}, Ser^{TCT}, Ser^{TCC}, Ser^{TCA} y Ser^{TCC}; y Thr^{ACC} tiene una clasificación de uso de codones superior pero una clasificación de preferencia de respuesta inmunitaria inferior con respecto a Thr^{ACG});

(ii) varios codones que Seed consideró que tenían una clasificación de uso de codones inferior en células de mamífero con respecto a al menos otro codón sinónimo tienen en realidad una clasificación de preferencia de respuesta inmunitaria superior al codón sinónimo o a cada uno de los codones sinónimo (por ejemplo, Ala^{GCT} tiene una clasificación de uso de codones inferior pero una clasificación de preferencia de respuesta inmunitaria superior con respecto a Ala^{GCC}; Gly^{GGA} tiene una clasificación de uso de codones inferior pero una clasificación de preferencia de respuesta inmunitaria superior con respecto a Gly^{GGC} o Gly^{GGG}; Phe^{TTT} tiene una clasificación de uso de codones inferior pero una clasificación de preferencia de respuesta inmunitaria superior con respecto a Phe^{TTC}; Ser^{TCG} tiene una clasificación de uso de codones inferior pero una clasificación de preferencia de respuesta inmunitaria superior con respecto a Ser^{AGC} o Ser^{TCC}; Ser^{TCT} y Ser^{TCA} tienen una clasificación de uso de codones inferior pero una clasificación de preferencia de respuesta inmunitaria superior con respecto a Ser^{AGC}; y Thr^{ACG} tiene una clasificación de uso de codones inferior pero una clasificación de preferencia de respuesta inmunitaria superior con respecto a Thr^{ACC});

(iii) varios codones que Seed consideró que tenían una clasificación de uso de codones superior en células de mamífero con respecto a otro codón sinónimo tienen de hecho la misma clasificación de preferencia de respuesta inmunitaria que el otro codón sinónimo (por ejemplo, Gln^{CAG} tiene una clasificación de uso de codones superior pero la misma clasificación de preferencia de respuesta inmunitaria que Gln^{CAA}; His^{CAC} tiene una clasificación de uso de codones superior pero la misma clasificación de preferencia de respuesta inmunitaria que His^{CAT}; Leu^{CTG} tiene una clasificación de uso de codones superior pero la misma clasificación de preferencia de respuesta inmunitaria que Leu^{CTC}; Lys^{AAG} tiene una clasificación de uso de codones superior pero la misma clasificación de preferencia de respuesta inmunitaria que Lys^{AAA}; Val^{GTG} tiene una clasificación de uso de codones superior pero la misma clasificación de preferencia de respuesta inmunitaria que Val^{GTC}); y

(iv) varios codones que Seed consideró que tenían la misma clasificación de uso de codones en células de mamífero que al menos otro codón sinónimo tienen de hecho una clasificación de preferencia de respuesta inmunitaria diferente que el codón sinónimo o codones sinónimos entre sí (por ejemplo, Ala^{GTC} tiene la misma clasificación de uso de codones pero una clasificación de preferencia de respuesta inmunitaria superior con respecto a Ala^{GCA} y Ala^{GCG}; Arg^{CGA}, Arg^{CGT} y Arg^{AGA} tienen la misma clasificación de uso de codones pero una clasificación de preferencia de respuesta inmunitaria superior con respecto a Arg^{AGG} y Arg^{CGG}; Glu^{GAA} tiene la misma clasificación de uso de codones pero una clasificación de preferencia de respuesta inmunitaria superior con respecto a Glu^{GAG}; Gly^{GGA} tiene la misma clasificación de uso de codones pero una clasificación de preferencia de respuesta inmunitaria superior con respecto a Gly^{GGT}; Leu^{CTA} y Leu^{CTT} tienen la misma clasificación de uso de codones pero una clasificación de preferencia de respuesta inmunitaria superior con respecto a Leu^{TTG} y Leu^{TTA}; Pro^{CCT} tiene la misma clasificación de uso de codones pero una clasificación de preferencia de respuesta inmunitaria superior con respecto a Pro^{CCA} o Pro^{CCG}; Ser^{TCG} tiene la misma clasificación de uso de codones pero una clasificación de preferencia de respuesta inmunitaria superior con respecto a uno cualquiera de Ser^{TCT}, Ser^{TCA} y Ser^{AGT}; Ser^{TCT} y Ser^{TCA} tienen la misma clasificación de uso de codones pero una clasificación de preferencia de respuesta inmunitaria superior con respecto a Ser^{AGT}; Thr^{ACG} tiene la misma clasificación de uso de codones pero una clasificación de preferencia de respuesta inmunitaria superior con respecto a un cualquiera de Thr^{ACA} y Thr^{ACT}; Thr^{ACG} tiene la misma clasificación de uso de codones pero una clasificación de preferencia de respuesta inmunitaria superior con respecto a Thr^{ACT}; Val^{GTT} tiene la misma clasificación de uso de codones pero una clasificación de preferencia de respuesta inmunitaria superior con respecto a Val^{GTA}).

En consecuencia, la presente divulgación permite por primera vez la modulación de una respuesta inmunitaria a un antígeno diana en un mamífero a partir de un polinucleótido que codifica un polipéptido que corresponde a al menos una porción del antígeno diana mediante el reemplazo de al menos un codón del polinucleótido por un codón sinónimo que tiene una preferencia superior o inferior para la producción de una respuesta inmunitaria con respecto al codón que reemplaza. En algunas realizaciones, por tanto, la presente invención abarca métodos de construcción de un polinucleótido sintético a partir del cual se puede producir un polipéptido para conferir una respuesta inmunitaria potenciada o más fuerte que una conferida por un polinucleótido parental que codifica el mismo polipéptido. Estos métodos en general comprenden seleccionar de la TABLA 1 un codón (con frecuencia

denominado en el presente documento arbitrariamente un "primer codón") del polinucleótido parental para el reemplazo por un codón sinónimo, en los que el codón sinónimo se selecciona basándose en que presenta una preferencia de respuesta inmunitaria superior con respecto al primer codón, y reemplazar el primer codón por el codón sinónimo para construir el polinucleótido sintético. Se hacen selecciones ilustrativas de los codones primero y sinónimo de acuerdo con la Tabla 2.

En algunas realizaciones, la selección de los codones primero y sinónimo se realiza de acuerdo con la TABLA 3, que es la misma que la Tabla 2 con la excepción de que excluye las selecciones basadas en las clasificaciones de uso de codones desvelada por Seed. En ejemplos ilustrativos de este tipo, la selección de un segundo codón (y posteriores codones si se desea) para el reemplazo por un codón sinónimo se hace de acuerdo con la TABLA 4.

Cuando se clasifican codones sinónimos en tres clasificaciones (clasificaciones 'alta', 'intermedia' y 'baja') en función de su clasificación de preferencia de respuesta inmunitaria (por ejemplo, los codones sinónimos para Ala, Ile, Leu, Pro, Ser, Thr y Val) se prefiere que el codón sinónimo que se selecciona sea un codón de clasificación alta cuando el primer codón es un codón de clasificación baja. Sin embargo, esto no es esencial y el codón sinónimo puede seleccionarse entre codones de clasificación intermedia. En el caso de dos o más codones sinónimos que tengan preferencias de respuesta inmunitaria similares, se apreciará que puede usarse cualquiera de estos codones para reemplazar el primer codón.

En otras realizaciones, la divulgación proporciona métodos de construcción de un polinucleótido sintético a partir del cual un se puede producir un polipéptido para conferir una respuesta inmunitaria reducida o más débil que una conferida por un polinucleótido parental que codifica el mismo polipéptido. Estos métodos en general comprenden seleccionar de la TABLA 1 un primer codón del polinucleótido parental para el reemplazo por un codón sinónimo, en los que el codón sinónimo se selecciona basándose en que presenta una preferencia de respuesta inmunitaria inferior con respecto al primer codón y reemplazar el primer codón con el codón sinónimo para construir el polinucleótido sintético. Se hacen selecciones ilustrativas de los codones primero y sinónimo de acuerdo con la TABLA 5.

En algunas realizaciones, la selección de los codones primero y sinónimo se hace de acuerdo con la TABLA 6, que es la misma que la TABLA 5 con la excepción de que excluye las selecciones basadas en las clasificaciones de uso de codones desveladas por Seed. En ejemplos ilustrativos de este tipo, la selección de un segundo codón (y codones posteriores si se desea) para el reemplazo por un codón sinónimo se hace de acuerdo con la TABLA 7.

Cuando se clasifican codones sinónimos en las tres clasificaciones indicadas anteriormente, se prefiere que el codón sinónimo que se selecciona sea un codón de clasificación baja cuando el primer codón es un codón de clasificación alta, pero esto no es esencial y, por tanto, el codón sinónimo puede seleccionarse entre codones de clasificación intermedia si se desea.

En general, la diferencia en la potencia de la respuesta inmunitaria producida en el mamífero a partir del polinucleótido sintético en comparación con la producida a partir del polinucleótido parental depende del número de codones primeros/segundos que son reemplazados por codones sinónimos y de la diferencia en la clasificación de preferencia de respuesta inmunitaria entre los codones primero/segundo y los codones sinónimos. Dicho de otro modo, a menor número de dichos reemplazos, y/o cuanto menor sea la diferencia en la clasificación de preferencia de respuesta inmunitaria entre los codones sinónimos y los codones primeros/segundos, menor será la diferencia entre la respuesta inmunitaria producida por el polinucleótido sintético y la producida por el polinucleótido parental. Por el contrario, a mayor número de dichos reemplazos, y/o cuanto mayor la diferencia en la clasificación de preferencia de respuesta inmunitaria entre los codones sinónimos y los codones primeros/segundos, mayor será la diferencia entre la respuesta inmunitaria producida por el polinucleótido sintético y la producida por el polinucleótido parental.

Es preferible pero no necesario reemplazar todos los codones del polinucleótido parental por codones sinónimos que tengan diferentes clasificaciones de preferencia de respuesta inmunitaria (por ejemplo, superior o inferior) con respecto a los codones primeros/segundos. Pueden lograrse cambios en la respuesta inmunitaria conferida incluso con un reemplazo parcial. En general, la etapa de reemplazo afecta al menos a aproximadamente el 5 %, el 10 %, el 15 %, el 20 %, el 25 %, el 30 %, por lo general al menos a aproximadamente el 35 %, el 40 %, el 50 % y normalmente al menos a aproximadamente el 60 %, el 70 %, el 80 %, el 90 %, el 95 %, el 96 %, el 97 %, el 98 %, el 99 % o más de los codones primeros/segundos del polinucleótido parental. En realizaciones en las que se necesita una respuesta inmunitaria más fuerte o potenciada, en general es deseable reemplazar algunos, preferentemente la mayoría y más preferentemente todos, los codones de clasificación baja en un polinucleótido parental con codones sinónimos que son codones de clasificación intermedia o preferentemente de clasificación alta. Normalmente, el reemplazo de codones de clasificación baja por codones de clasificación intermedia o alta dará como resultado un aumento de la potencia de la respuesta inmunitaria a partir del polinucleótido sintético construido de este modo, en comparación con el producido a partir del polinucleótido parental en las mismas condiciones. Sin embargo, con frecuencia es deseable reemplazar algunos, preferentemente la mayoría y más preferentemente todos, los codones de clasificaciones intermedia en el polinucleótido parental por codones de clasificación alta, si se desean respuestas inmunitarias más fuertes o más potenciadas.

Por el contrario, en algunas realizaciones en las que se requiere una respuesta inmunitaria más débil o reducida, en general es deseable reemplazar algunos, preferentemente la mayoría y más preferentemente todos, los codones de clasificación alta en un polinucleótido parental por codones sinónimos que sean codones de clasificación intermedia o preferentemente baja. Normalmente, el reemplazo de codones de clasificación alta por codones de clasificación intermedia o baja dará como resultado una disminución sustancial en la potencia de la respuesta inmunitaria a partir del polinucleótido sintético construido de este modo, en comparación con la producida a partir del polinucleótido parental en las mismas condiciones. En realizaciones específicas en las que se desea conferir una respuesta inmunitaria más débil o más reducida, en general es deseable reemplazar algunos, preferentemente la mayoría y más preferentemente todos, los codones de clasificación intermedia en el polinucleótido parental por codones de clasificación baja.

En ejemplos ilustrativos que requieren una respuesta inmunitaria más fuerte o potenciada, el número y la diferencia en la clasificación de preferencia de respuesta inmunitaria entre los codones primeros/segundos y los codones sinónimo se seleccionan de manera que la respuesta inmunitaria conferida por el polinucleótido sintético sea al menos aproximadamente el 110 %, el 150 %, el 200 %, el 250 %, el 300 %, el 350 %, el 400 %, el 450 %, el 500 %, el 600 %, el 700 %, el 800 %, el 900 %, el 1000 % o más, de la respuesta inmunitaria conferida por el polinucleótido parental en las mismas condiciones. Por el contrario, en algunas realizaciones que requieren una respuesta inmunitaria menor o más débil, el número y la diferencia en la clasificación de preferencia de respuesta inmunitaria entre los codones primeros/segundos y los codones sinónimo se seleccionan de manera que la respuesta inmunitaria conferida por el polinucleótido sintético no sea superior a aproximadamente el 90 %, el 80 %, el 70 %, el 60 %, el 50 %, el 40 %, el 30 %, el 20 %, el 10 %, el 5 % o menos de la respuesta inmunitaria conferida por el polinucleótido parental en las mismas condiciones.

7. Modulación de respuestas inmunitarias en mamíferos mediante la expresión de polinucleótidos que codifican ARN de transferencia isoaceptor

Es posible aprovechar las clasificaciones de preferencia de respuesta inmunitaria de los codones analizados en la Sección 6 para modular una respuesta inmunitaria a un antígeno diana cambiando el nivel de iso-ARNt en la población celular que es la diana de la inmunización. En consecuencia, la divulgación también presenta métodos de potenciación de la cualidad de una respuesta inmunitaria a un antígeno diana en un mamífero, en los que la respuesta es conferida por la expresión de un primer polinucleótido que codifica un polipéptido que corresponde a al menos una porción del antígeno diana. Estos métodos en general comprenden: introducir en el mamífero una primera construcción de ácido nucleico que comprende el primer polinucleótido en conexión operable con una secuencia reguladora. Después se introduce una segunda construcción de ácido nucleico en el mamífero, que comprende un segundo polinucleótido que está conectado operativamente a una secuencia reguladora y que codifica un iso-ARNt que corresponde a un codón de preferencia inmunitaria baja del primer polinucleótido.

En la práctica, por tanto, se introduce un iso-ARNt en el mamífero mediante la segunda construcción de ácido nucleico cuando el iso-ARNt corresponde a un codón de preferencia de respuesta inmunitaria baja en el primer polinucleótido, que se selecciona convenientemente entre el grupo que consiste en Ala^{GCA}, Ala^{GCG}, Ala^{GCC}, Arg^{AGG}, Arg^{CGG}, Asn^{AAT}, Asp^{GAT}, Cys^{TGT}, Glu^{GAG}, Gly^{GGG}, Gly^{GGT}, Gly^{GGC}, Ile^{ATA}, Ile^{ATT}, Leu^{TTG}, Leu^{TTA}, Leu^{CTA}, Leu^{CTT}, Phe^{TTC}, Pro^{CCA}, Pro^{CCG}, Pro^{CCT}, Ser^{AGC}, Ser^{AGT}, Ser^{TCT}, Ser^{TCA}, Ser^{TCC}, Thr^{ACA}, Thr^{ACT}, Tyr^{TAT}, Val^{GTA} y Val^{GTT}. En realizaciones específicas, los iso-ARNt suministrados son específicos para codones que tienen codones de preferencia de respuesta inmunitaria 'baja', que pueden seleccionarse entre el grupo que consiste en Ala^{GCA}, Ala^{GCG}, Arg^{AGG}, Arg^{CGG}, Asn^{AAT}, Asp^{GAT}, Cys^{TGT}, Glu^{GAG}, Gly^{GGG}, Gly^{GGT}, Gly^{GGC}, Ile^{ATA}, Leu^{TTG}, Leu^{TTA}, Phe^{TTC}, Pro^{CCA}, Pro^{CCG}, Ser^{AGC}, Ser^{AGT}, Thr^{ACT}, Tyr^{TAT} y Val^{GTA}. La primera construcción (es decir, la construcción que expresa antígeno) y la segunda construcción (es decir, la construcción que expresa iso-ARNt) pueden introducirse simultáneamente o secuencialmente (en cualquier orden) y pueden introducirse en el mismo o diferentes sitios. En algunas realizaciones, las construcciones primera y segunda están contenidas en vectores separados. En otras realizaciones, están contenidas en un solo vector. Si se desea, pueden introducirse dos o más segundas construcciones expresando, cada una, un iso-ARNt diferente que corresponde a un codón de preferencia baja del primer polinucleótido. Las construcciones de ácido nucleico primera y segunda pueden construirse y administrarse de forma concurrente o simultáneamente a un mamífero de acuerdo con cualquier procedimiento adecuado, cuyos ejemplos ilustrativos se analizan a continuación para las construcciones químicas.

En algunas realizaciones, se administra una pluralidad de diferentes construcciones que expresan iso-ARNt (por ejemplo, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 o más) simultáneamente o contemporáneamente con la construcción que expresa antígeno, en las que las construcciones que expresan iso-ARNt individuales expresan un iso-ARNt diferente que otras construcciones que expresan iso-ARNt.

8. Antígenos

Son antígenos diana útiles en la presente invención normalmente moléculas proteínicas, cuyos ejemplos representativos incluyen polipéptidos y péptidos. Los antígenos diana pueden seleccionarse entre antígenos endógenos producidos por un hospedador o antígenos exógenos que son extraños para el hospedador. Los antígenos endógenos adecuados incluyen, pero no se limitan a, antígenos de cáncer o tumorales. Los ejemplos no

limitantes de antígenos de cáncer o tumorales incluyen antígenos de un cáncer o tumor seleccionados entre protooncogén ABL1, cánceres relacionados con el SIDA, neuroma acústico, leucemia linfocítica aguda, leucemia mieloide aguda, carcinoma adenoquístico, cáncer adrenocortical, metaplasia mieloide agnógena, alopecia, sarcoma alveolar de partes blandas, cáncer anal, angiosarcoma, anemia aplásica, astrocitoma, ataxia-telangiectasia, carcinoma de células basales (piel), cáncer de vejiga, cánceres de hueso, cáncer de intestino, glioma del tronco cerebral, el cerebro y tumores del SNC, cáncer de mama, tumores del SNC, tumores carcinoides, cáncer de cuello uterino, tumores cerebrales infantiles, cáncer infantil, leucemia infantil, sarcoma infantil del tejido blando, condrosarcoma, coriocarcinoma, leucemia linfocítica crónica, leucemia mieloide crónica, cáncer colorrectal, linfoma cutáneo de linfocitos T, dermatofibrosarcoma protuberante, tumores desmoplásicos de células redondas pequeñas, carcinoma ductal, cáncer endocrino, cáncer de endometrio, ependimoma, cáncer esofágico, sarcoma de Ewing, cáncer extrahepático de las vías biliares, cáncer de ojo, Ojo: melanoma, retinoblastoma, cáncer de Trompas de Falopio, anemia de Fanconi, fibrosarcoma, cáncer de vesícula biliar, cáncer de estómago, cánceres gastrointestinales, tumor-carcinoide gastrointestinal, cánceres genitourinarios, tumores de células germinales, enfermedad trofoblástica gestacional, glioma, cánceres ginecológicos, neoplasias hemáticas, leucemia de células pilosas, cáncer de cabeza y cuello, cáncer hepatocelular, cáncer de mama hereditario, histiocitosis, enfermedad de Hodgkin, virus del papiloma humano, mola hidatidiforme, hipercalcemia, cáncer de hipofaringe, melanoma intraocular, cáncer de células de los islotes, sarcoma de Kaposi, cáncer de riñón, histiocitosis de células de Langerhans, cáncer de laringe, leiomioma, leucemia, síndrome de Li-Fraumeni, cáncer de labio, liposarcoma, cáncer de hígado, cáncer de pulmón, linfedema, linfoma, linfoma de Hodgkin, linfoma no Hodgkin, cáncer de mama masculino, tumor rabdoide maligno de riñón, meduloblastoma, melanoma, cáncer de células de Merkel, mesotelioma, cáncer metastásico, cáncer de boca, neoplasia endocrina múltiple, micosis fungoide, síndromes mielodisplásicos, mieloma, trastornos mieloproliferativos, cáncer nasal, cáncer de la nasofaringe, nefroblastoma, neuroblastoma, neurofibromatosis, síndrome de rotura de Nijmegen, cáncer de piel no melanoma, cáncer de pulmón no microcítico (CPNM), cánceres oculares, cáncer de esófago, cáncer de la cavidad oral, cáncer de la orofaringe, osteosarcoma, cáncer de ovario por ostomía, cáncer de páncreas, cáncer paranasal, cáncer de paratiroides, cáncer de la glándula parótida, cáncer de pene, tumores periféricos-neuroectodérmicos, cáncer de hipófisis, policitemia vera, cáncer de próstata, cánceres raros y trastornos asociados, carcinoma de células renales, retinoblastoma, rabdomiosarcoma, síndrome de Rothmund-Thomson, cáncer de la glándula salival, sarcoma, schwannoma, síndrome de Sezary, cáncer de piel, cáncer de pulmón microcítico (CPM), cáncer de intestino delgado, sarcoma de tejidos blandos, tumores de la médula, carcinoma de células escamosas (piel), cáncer de estómago, sarcoma sinovial, cáncer testicular, cáncer del timo, cáncer de tiroides, cáncer de células de transición (vejiga), cáncer de células de transición (renal-pelvis/- uréter), cáncer trofoblástico, cáncer de uretra, cáncer del sistema urinario, uroplaquinas, sarcoma uterino, cáncer de útero, cáncer vaginal, cáncer de la vulva, macroglobulinemia de Waldenstrom, tumor de Wilms. En ciertas realizaciones, el cáncer o tumor se refiere a melanoma. Los ejemplos ilustrativos de antígenos relacionados con el melanoma incluyen el antígeno de diferenciación de melanocitos (por ejemplo, gp100, MART, Melan-A/MART-1, TRP-1, Tyros, TRP2, MC1R, MUC1F, MUC1R o una combinación de los mismos) y los antígenos específicos de melanoma (por ejemplo, BAGE, GAGE-1, gp100In4, MAGE-1 (e.g., GenBank Accession No. X54156 and AA4943311), MAGE-3, MAGE4, PRAME, TRP2IN2, NYNSO1a, NYNSO1b, LAGE1, antígeno de melanoma p97 (por ejemplo, n.º de Acceso de GenBank M12154), proteína p5, gp75, antígeno oncofetal, gangliósidos GM2 y GD2, cdc27, p21ras, gp100^{Pmel117} o una combinación de los mismos. Otros antígenos específicos de tumor incluyen, pero no se limitan a: etv6, aml1, ciclofilina B (leucemia linfoblástica aguda); idioma de Ig (linfoma de células B); E-cadherina, α -catenina, β -catenina, γ -catenina, p120ctn (glioma); p21ras (cáncer de vejiga); p21ras (cáncer biliar); familia MUC, HER2/neu, c-erbB-2 (cáncer de mama); p53, p21ras (carcinoma cervical); p21ras, HER2/neu, c-erbB-2, familia MUC, proteína Cripto-1, proteína Pim-1 (carcinoma de colon); antígeno asociado al sistema colorrectal (CRC)-CO17-1A/GA733, APC (cáncer colorrectal); antígeno carcinoembrionario (ACE) (cáncer colorrectal; coriocarcinoma); ciclofilina B (cáncer de células epiteliales); HER2/neu, c-erbB-2, glucoproteína ga733 (cáncer gástrico); α -fetoproteína (cáncer hepatocelular); Imp-1, EBNA-1 (linfoma de Hodgkin); ACE, MAGE-3, NY-ESO-1 (cáncer de pulmón); ciclofilina B (leucemia derivado de células linfoides); familia MUC, p21ras (mieloma); HER2/neu, c-erbB-2 (carcinoma de pulmón no microcítico); Imp-1, EBNA-1 (cáncer nasofaríngeo); familia MUC; HER2/neu, c-erbB-2, MAGE-A4, NY-ESO-1 (cáncer de ovario); Antígeno prostático específico (PSA) y su epítopos antígenicos PSA-1, PSA-2 y PSA-3, PSMA, HER2/neu, c-erbB-2, glucoproteína ga733 (cáncer de próstata); HER2/neu, c-erbB-2 (cáncer renal); productos víricos tales como proteínas de papilomavirus humano (cánceres de células escamosas del cuello uterino y el esófago); NY-ESO-1 (cáncer testicular); y epítopos de VLTH-1 (leucemia de linfocitos T).

Se seleccionan convenientemente antígenos extraños o exógenos entre antígenos de organismos patógenos. Los organismos patógenos de ejemplo incluyen, pero no se limitan a, virus, bacterias, parásitos hongos, algas y protozoos y amebas. Los virus ilustrativos incluyen virus responsables de enfermedades incluyendo, pero no limitadas a, sarampión, paperas, rubéola, poliomieltis, hepatitis A, B (por ejemplo, N.º de Acceso de GenBank E02707) y C (por ejemplo, N.º de Acceso de GenBank E06890), así como otros virus de hepatitis, gripe, adenovirus (por ejemplo, tipos 4 y 7), rabia (por ejemplo, N.º de Acceso de GenBank M34678), fiebre amarilla, virus de Epstein-Barr y otros herpesvirus, tales como el virus del papiloma, virus Ebola, virus de la gripe, encefalitis japonesa (por ejemplo, N.º de Acceso de GenBank E07883), dengue (por ejemplo, N.º de Acceso de GenBank M24444), hantavirus, virus Sendai, virus respiratorio sincitial, ortomixovirus, virus de la estomatitis vesicular, virus visna, citomegalovirus y virus de la inmunodeficiencia humana (VIH) (por ejemplo, N.º de Acceso de GenBank U18552). Es útil cualquier antígeno adecuado derivado de dichos virus en la práctica de la presente invención. Por ejemplo, los

antígenos retrovíricos ilustrativos derivados de VIH incluyen, pero no se limitan a, antígenos tales como productos génicos de los genes gag, pol y env, la proteína Nef, transcriptasa inversa y otros componentes del VIH. Los ejemplos ilustrativos de antígenos víricos de hepatitis incluyen, pero no se limitan a, antígenos tales como las proteínas S, M y L del virus de la hepatitis B, el antígeno pre-S del virus de la hepatitis B, y otras hepatitis, por ejemplo, las hepatitis A, B y C, componentes víricos tales como el ARN vírico de la hepatitis C. Los ejemplos ilustrativos de antígenos víricos de gripe incluyen; pero no se limitan a, antígenos tales como la hemaglutinina y la neuraminidasa y otros componentes víricos de la gripe. Los ejemplos ilustrativos de antígenos víricos del sarampión incluyen, pero no se limitan a, antígenos tales como la proteína de fusión del virus del sarampión y otros componentes del virus del sarampión. Los ejemplos ilustrativos de antígenos víricos de la rubéola incluyen, pero no se limitan a, antígenos tales como las proteínas E1 y E2 y otros componentes del virus de la rubéola; antígenos de rotavirus como VP7sc y otros componentes del rotavirus. Los ejemplos ilustrativos de los antígenos de citomegalovirus incluyen, pero no se limitan a, antígenos tales como la glucoproteína de envoltura B y otros componentes antigénicos de citomegalovirus. Los ejemplos no limitantes de antígenos víricos del virus respiratorio sincitial incluyen antígenos tales como la proteína de fusión del VRS, la proteína M2 y otros componentes de antígenos del virus respiratorio sincitial. Los ejemplos ilustrativos de antígenos víricos del herpes simple incluyen, pero no se limitan a, antígenos tales como proteínas inmediatas tempranas, glucoproteína D y otros componentes antigénicos víricos del herpes simple. Los ejemplos no limitantes de antígenos víricos de la varicela zóster incluyen antígenos tales como 9PI, gpII y otros componentes antigénicos víricos de la varicela zóster. Los ejemplos no limitantes de antígenos víricos de encefalitis japonesa incluyen antígenos tales como las proteínas E, ME, ME-NS 1, NS 1, NS 1-NS2A, E 80 % y otros componentes antigénicos víricos de encefalitis japonesa. Los ejemplos representativos de antígenos víricos de la rabia incluyen, pero no se limitan a, antígenos tales como la glucoproteína de la rabia, la nucleoproteína de la rabia y otros componentes antigénicos víricos de la rabia. Los ejemplos ilustrativos de antígenos de papilomavirus incluyen, pero no se limitan a, las proteínas de la cápside L1 y L2, así como los antígenos E6/E7 asociados a cánceres cervicales, Véase *Fundamental Virology*, segunda edición, eds. Fields, B.N. y Knipe, D.M., 1991, Raven Press, Nueva York, para ejemplos adicionales de antígenos víricos.

Los ejemplos ilustrativos de hongos incluyen *Acremonium spp.*, *Aspergillus spp.*, *Basidiobolus spp.*, *Bipolaris spp.*, *Blastomyces dermatidis*, *Candida spp.*, *Cladophialophora carrionii*, *Coccidioides immitis*, *Conidiobolus spp.*, *Cryptococcus spp.*, *Curvularia spp.*, *Epidermophyton spp.*, *Exophiala jeanselmei*, *Exserohilum spp.*, *Fonsecaea compacta*, *Fonsecaea pedrosoi*, *Fusarium oxysporum*, *Fusarium solani*, *Geotrichum candidum*, *Histoplasma capsulatum* var. *capsulatum*, *Histoplasma capsulatum* var. *duboisii*, *Hortaea werneckii*, *Lacazia loboi*, *Lasiodiplodia theobromae*, *Leptosphaeria senegalensis*, *Madurella grisea*, *Madurella mycetomatis*, *Malassezia furfur*, *Microsporum spp.*, *Neotestudina rosatii*, *Onychocola canadensis*, *Paracoccidioides brasiliensis*, *Phialophora verrucosa*, *Piedraia hortae*, *Piedra iahortae*, *Pityriasis versicolor*, *Pseudallescheria boydii*, *Pyrenochaeta romeroi*, *Rhizopus arrhizus*, *Scopulariopsis brevicaulis*, *Scytalidium dimidiatum*, *Sporothrix schenckii*, *Trichophyton spp.*, *Trichosporon spp.*, *Zygomycete fungi*, *Absidia corymbifera*, *Rhizomucor pusillus* y *Rhizopus arrhizus*. Por tanto, los antígenos fúngicos representativos que pueden usarse en las composiciones y métodos de la presente invención incluyen, pero no se limitan a, componentes antigénicos fúngicos de *Candida*; antígenos fúngicos de *Histoplasma* tales como la proteína de choque térmico 60 (HSP60) y otros componentes antigénicos fúngicos de *Histoplasma*; antígenos fúngicos criptocócicos tales como polisacáridos capsulares y otros componentes antigénicos fúngicos criptocócicos; antígenos fúngicos de *Coccidioides* tales como antígenos de esférulas y otros componentes antigénicos fúngicos de *Coccidioides*; y antígenos fúngicos de la tiña tales como la tricofitina y otros componentes antigénicos fúngicos de *Coccidioides*.

Los ejemplos ilustrativos de bacterias incluyen bacterias que son responsables de enfermedades incluyendo, pero no limitadas a, difteria (por ejemplo, *Corynebacterium diphtheria*), tosferina (por ejemplo, *Bordetella pertussis*, N.º de Acceso de GenBank M35274), tétanos (por ejemplo, *Clostridium tetani*, N.º de Acceso de GenBank M64353), tuberculosis (por ejemplo, *Mycobacterium tuberculosis*), neumonías bacterianas (por ejemplo, *Haemophilus influenzae*), cólera (por ejemplo, *Vibrio cholerae*), carbunco (por ejemplo, *Bacillus anthracis*), fiebre tifoidea, peste, shigelosis (por ejemplo, *Shigella dysenteriae*), botulismo (por ejemplo, *Clostridium botulinum*), salmonelosis (por ejemplo, N.º de Acceso GenBank L03833), úlceras pépticas (por ejemplo, *Helicobacter pylori*), la enfermedad del legionario, la enfermedad de Lyme (por ejemplo, N.º de Acceso GenBank U59487). Otras bacterias patógenas incluyen *Escherichia coli*, *Clostridium perfringens*, *Pseudomonas aeruginosa*, *Staphylococcus aureus* y *Streptococcus pyogenes*. Por tanto, los antígenos bacterianos que pueden usarse en las composiciones y métodos de la invención incluyen, pero no se limitan a: antígenos bacterianos de la tosferina tales como la toxina pertúsica, hemaglutinina filamentosa, pertactina, F M2, FIM3, adenilato ciclasa y otros componentes antigénicos bacterianos de la tosferina; antígenos bacterianos de la difteria, tales como la toxina o el toxoide diftérico y otros componentes antigénicos bacterianos de la difteria; antígenos bacterianos del tétanos tales como la toxina o el toxoide tetánicos y otros componentes antigénicos bacterianos del tétanos, antígenos bacterianos de estreptococos, tales como las proteínas M y otros componentes antigénicos bacterianos de estreptococos; antígenos bacterianos de bacilos gramnegativos tales como lipopolisacáridos y otros componentes antigénicos de bacterias gramnegativas; antígenos bacterianos de *Mycobacterium tuberculosis*, tales como ácidos micólicos, la proteína de choque térmico 65 (HSP65), la proteína secretada mayor de 30 kDa, antígeno 85A y otros componentes antigénicos de micobacterias; componentes antigénicos bacterianos de *Helicobacter pylori*, antígenos bacterianos neumocócicos tales como neumolisina, polisacáridos capsulares de neumococo y otros componentes antigénicos bacterianos de neumococo; antígenos bacterianos de *Haemophilus influenza* tales como polisacáridos capsulares y otros componentes

antigénicos bacterianos de *Haemophilus influenza*; antígenos bacterianos del carbunco tales como el antígeno protector del carbunco y otros componentes antigénicos bacterianos del carbunco; antígenos bacterianos de las rickettsias, tales como rOmpA y otros componentes antigénicos bacterianos de las rickettsias. También se incluye con los antígenos bacterianos que se describen en el presente documento cualquier otro antígeno bacteriano, micobacteriano, de micoplasmas, de rickettsias o de clamidias.

Los ejemplos ilustrativos de protozoos incluyen protozoos que son responsables de enfermedades incluyendo, pero no limitadas a, malaria (por ejemplo, N.º de Acceso de GenBank X53832), anquilostomiasis, oncocercosis (por ejemplo, N.º de Acceso de GenBank M27807), esquistosomiasis (por ejemplo, N.º de Acceso de GenBank LOS 198), toxoplasmosis, tripanosomiasis, leishmaniasis, giardiasis (N.º de Acceso de GenBank M33641), amebiasis, filariasis (por ejemplo, N.º de Acceso de GenBank J03266), borreliosis y triquinosis. Por tanto, los antígenos de protozoos que pueden usarse en las composiciones y métodos de la invención incluyen, pero no se limitan a: antígenos de *Plasmodium falciparum*, tales como antígenos de superficie de merozoítos, antígenos de superficie de esporozoítos, antígenos de circunsporozoítos, antígenos de superficie de gametocitos/gametos, antígeno de la etapa en sangre pf 155/RESA y otros componentes antigénicos de plasmodios; antígenos de *Toxoplasma* tales como SAG-1, p30 y otros componentes antigénicos de toxoplasma; antígenos de *Schistosoma* tales como glutatión-S-transferasa, paramiosina y otros componentes antigénicos de *Schistosoma*; *Leishmania* mayor y otros antígenos de *Leishmania* tales como gp63, lipofosfoglicano y su proteína asociada y otros componentes antigénicos de *Leishmania*; y antígenos de *Trypanosoma cruzi* tales como el antígeno de 75-77 kDa, el antígeno de 56 kDa y otros componentes antigénicos de *Trypanosoma*.

La presente divulgación también incluye componentes de toxinas como antígenos, cuyos ejemplos ilustrativos incluyen enterotoxinas estafilocócicas, toxina del síndrome de choque tóxico; antígenos retrovíricos (por ejemplo, antígenos derivados del VIH), antígenos estreptocócicos, enterotoxina estafilocócica A (SEA), enterotoxina estafilocócica B (SEB), enterotoxina estafilocócica 1-3 (SE₁₋₃), enterotoxina estafilocócica-D (SED), enterotoxina estafilocócica E (SEE), así como toxinas derivadas de *Mycoplasma*, *Mycobacterium* y del virus del herpes.

9. Construcción de polinucleótidos sintéticos

El reemplazo de un codón por otro puede conseguirse usando métodos convencionales conocidos en la técnica. Por ejemplo, la modificación por codones de un polinucleótido parental puede efectuarse usando varias técnicas de mutagénesis conocidas incluyendo, por ejemplo, la mutagénesis dirigida a oligonucleótidos, la mutagénesis con oligonucleótidos degenerados y la mutagénesis específica de región. Se describen técnicas de mutagénesis *in vitro*, por ejemplo, en las Patentes de los EE.UU. N.º 4.184.917, 4.321.365 y 4.351.901 o en las secciones pertinentes de Ausubel, et al. (*CURRENT PROTOCOLS IN MOLECULAR BIOLOGY*, John Wiley & Sons, Inc., 1997) y de Sambrook, et al., (*MOLECULAR CLONING. A LABORATORY MANUAL*, Cold Spring Harbor Press, 1989). En lugar de mediante mutagénesis *in vitro*, el polinucleótido sintético puede sintetizarse *de novo* usando maquinaria fácilmente disponible como se describe, por ejemplo, en la Patente de los EE.UU. N.º 4.293.652. Sin embargo, hay que señalar que la presente invención no depende de y no se refiere a, ninguna técnica particular para construir el polinucleótido sintético.

El polinucleótido parental es convenientemente un gen natural. Sin embargo, es posible que el polinucleótido parental no sea de origen natural pero que se ha diseñado usando técnicas recombinantes. Pueden obtenerse polinucleótidos parentales a partir de cualquier fuente adecuada, tal como a partir de organismos eucariotas o procariotas, incluyendo, pero no limitados a mamíferos u otros animales y organismos patógenos tales como levaduras, bacterias, protozoos y virus.

La divulgación también contempla polinucleótidos sintéticos que codifican una o más porciones de un antígeno diana deseado. En algunas realizaciones, el polinucleótido sintético codifica al menos aproximadamente 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 40, 50, 60, 70, 80, 90, 100, 120, 150, 300, 400, 500, 600, 700, 800, 900 o 1000 o incluso al menos aproximadamente 2000, 3000, 4000 o 5000 restos de aminoácidos contiguos, o casi hasta el número total de aminoácidos presentes en un antígeno diana de longitud completa. En algunas realizaciones, el polinucleótido sintético codifica una pluralidad de porciones del antígeno diana, en el que las porciones son iguales o diferentes. En ejemplos ilustrativos de este tipo, el polinucleótido sintético codifica una proteína de fusión de múltiples epítomos. Una serie de factores pueden influir en la elección del tamaño de la porción. Por ejemplo, el tamaño de las porciones individuales codificadas por el polinucleótido sintético puede elegirse de manera que incluya o se corresponda con el tamaño de epítomos de linfocitos T y/o epítomos de células B y sus requisitos de procesamiento. Los practicantes en la materia reconocerán que los epítomos de linfocitos T restringidos de clase I normalmente tienen entre 8 y 10 restos de aminoácidos de longitud y, si se colocan junto a restos flanqueantes no naturales, dichos epítomos en general pueden necesitar de 2 a 3 restos de aminoácidos flanqueantes naturales para asegurar que se procesen y se presenten de manera eficiente. Los epítomos de linfocitos T restringidos de clase II por lo general varían entre 12 y 25 restos de aminoácidos de longitud y pueden no necesitar restos flanqueantes naturales para el procesamiento proteolítico eficaz, aunque se cree que los restos flanqueantes naturales pueden desempeñar un papel. Otra característica importante de los epítomos restringidos de clase II es que en general contienen un núcleo de 9-10 restos de aminoácidos en el medio que se unen específicamente a moléculas del CMH de clase II con secuencias flanqueantes a ambos lados de este núcleo.

estabilizando la unión mediante la asociación con estructuras conservadas a cada lado de antígenos del CMH de clase II de una forma independiente de la secuencia. De este modo, la región funcional de los epítomos restringidos de clase II es normalmente inferior a aproximadamente 15 restos de aminoácidos de longitud. El tamaño de los epítomos de células B lineales y los factores que afectan a su procesamiento, como los epítomos restringidos de clase II, son muy variables, aunque dichos epítomos con frecuencia son inferiores en tamaño a 15 restos de aminoácidos. A partir de lo anterior, es ventajoso, pero no esencial, que el tamaño de las porciones individuales del antígeno diana sea de al menos 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 25, 30 restos de aminoácidos. Convenientemente, el tamaño de las porciones individuales no es superior a aproximadamente 500, 200, 100, 80, 60, 50, 40 restos de aminoácidos. En ciertas realizaciones ventajosas, el tamaño de las porciones individuales es suficiente para la presentación por una célula presentadora de antígeno de un epítomo de linfocito T y/o un epítomo de célula B contenidos dentro del péptido.

Como se apreciará por los expertos en la materia, en general no es necesario inmunizar con un polipéptido que comparta exactamente la misma secuencia de aminoácidos con el antígeno diana para producir una respuesta inmunitaria a ese antígeno. En algunas realizaciones, por tanto, el polipéptido codificado por el polinucleótido sintético es deseablemente una variante de al menos una porción del antígeno diana. Los polipéptidos "variantes" incluyen proteínas derivadas del antígeno diana por delección (denominada truncamiento) o adición de uno o más aminoácidos al extremo N terminal y/o C terminal del antígeno diana; la delección o adición de uno o más aminoácidos en uno o más sitios en el antígeno diana; o la sustitución de uno o más aminoácidos en uno o más sitios en el antígeno diana. Los polipéptidos variantes tendrán al menos un 40 %, un 50 %, un 60 %, un 70 %, en general al menos un 75 %, un 80 %, un 85 %, normalmente al menos aproximadamente de un 90 % a un 95 % o más y más normalmente al menos aproximadamente un 96 %, un 97 %, un 98 %, un 99 % o más de similitud o identidad de secuencia con la secuencia de aminoácidos del antígeno diana o una porción del mismo como se determina mediante programas de alineación de secuencia descritos en otra parte del presente documento usando parámetros por defecto. Una variante de un antígeno diana puede diferir de ese antígeno en general en tanto como 1000, 500, 400, 300, 200, 100, 50 o 20 restos de aminoácidos o, convenientemente, en tan poco como 1-15 restos de aminoácidos, tan poco como 1-10, tal como 6-10, tan poco como 5, tan poco como 4, 3, 2 o incluso 1 resto de aminoácido.

Los polipéptidos variantes que corresponden a al menos una porción de un antígeno diana pueden contener sustituciones conservadoras de aminoácidos en diversas localizaciones a lo largo de su secuencia, en comparación con la secuencia de aminoácidos del antígeno diana. Una "sustitución de aminoácidos conservadora" es una en la que el resto de aminoácido se reemplaza por un resto de aminoácido que tiene una cadena lateral similar. Las familias de restos de aminoácidos que tienen cadenas laterales similares se han definido en la técnica, que en general pueden subclasificarse como se indica a continuación:

Ácidos: El resto tiene una carga negativa debido a la pérdida de iones de H a pH fisiológico y el resto es atraído por la solución acuosa con el fin de buscar las posiciones superficiales en la conformación de un péptido en el que está contenido cuando el péptido está en medio acuoso a pH fisiológico. Los aminoácidos que tienen una cadena lateral ácida incluyen el ácido glutámico y el ácido aspártico.

Básicos: El resto tiene una carga positiva debido a la asociación con el ion H a pH fisiológico o a una o dos unidades de pH del mismo (por ejemplo, histidina) y el resto es atraído por la solución acuosa con el fin de buscar las posiciones superficiales en la conformación de un péptido en el que está contenido cuando el péptido está en medio acuoso a pH fisiológico. Los aminoácidos que tienen una cadena lateral básica incluyen arginina, lisina e histidina.

Cargados: Los restos están cargados a pH fisiológico y, por tanto, incluyen aminoácidos que tienen cadenas laterales ácidas o básicas (es decir, ácido glutámico, ácido aspártico, arginina, lisina e histidina).

Hidrófobos: Los restos no están cargados a pH fisiológico y el resto es repelido por la solución acuosa con el fin de buscar las posiciones internas en la conformación de un péptido en el que está contenido cuando el péptido está en medio acuoso. Los aminoácidos que tienen una cadena lateral hidrófoba incluyen tirosina, valina, isoleucina, leucina, metionina, fenilalanina y triptófano.

Neutros/polares: Los restos no están cargados a pH fisiológico, pero el resto no es repelido suficientemente por las soluciones acuosas de manera que buscaría posiciones internas en la conformación de un péptido en el que está contenido cuando el péptido está en medio acuoso. Los aminoácidos que tienen una cadena lateral neutra/polar incluyen asparagina, glutamina, cisteína, histidina, serina y treonina.

La presente descripción también caracteriza ciertos aminoácidos como "pequeños" puesto que sus cadenas laterales no son suficientemente grandes, incluso si faltan grupos polares, para conferir hidrofobia. Con la excepción de la prolina, los aminoácidos "pequeños" son aquellos con cuatro carbonos o menos cuando hay al menos un grupo polar en la cadena lateral y tres carbonos o menos cuando no. Los aminoácidos que tienen una cadena lateral pequeña incluyen glicina, serina, alanina y treonina. El aminoácido secundario codificada por un gen, prolina, es un caso especial debido a sus efectos conocidos sobre la conformación secundaria de cadenas de péptidos. La estructura de la prolina se diferencia de todos los demás aminoácidos de origen natural en que su cadena lateral

está unida al nitrógeno del grupo α -amino, así como al α -carbono. Varias matrices de similitud de aminoácidos (por ejemplo, la matriz PAM120 y la matriz PAM250 como se desvelan por ejemplo por Dayhoff et al. (1978) *A model of evolutionary change in proteins. Matrices for determining distance relationships*, en M. O. Dayhoff, (ed.), *Atlas of protein sequence and structure*, Vol 5, págs. 345-358, National Biomedical Research Foundation, Washington DC y por Gonnet et al., 1992, *Science* 256 (5062): 144301445), sin embargo, incluyen la prolina en el mismo grupo que la glicina, la serina, la alanina y la treonina. En consecuencia, para los fines de la presente invención, la prolina se clasifica como un aminoácido "pequeño".

El grado de atracción o repulsión necesario para la clasificación como polar o no polar es arbitrario y, por tanto, los aminoácidos se han clasificado como uno u otro. La mayoría de los aminoácidos no expresamente mencionados pueden clasificarse basándose en el comportamiento conocido.

Los restos de aminoácidos pueden subclasificarse adicionalmente como cíclicos o no cíclicos y aromáticos o no aromáticos, clasificaciones autoexplicativas con respecto a los grupos sustituyentes de la cadena lateral de los restos, y como pequeños o grandes. El resto se considera pequeño si contiene un total de cuatro átomos de carbono o menos, incluido el carbono del carboxilo, a condición de que esté presente un sustituyente polar adicional; tres o menos si no lo está. Los residuos pequeños son siempre, por supuesto, no aromáticos. Dependiendo de sus propiedades estructurales, los restos de aminoácidos pueden pertenecer a dos o más clases. Para los aminoácidos de proteínas de origen natural, la subclasificación de acuerdo con este esquema se presenta en la Tabla 10.

TABLA 10

Resto original	Sustituciones de ejemplo
Ala	Ser
Arg	Lys
Asn	Gln, His
Asp	Glu
Cys	Ser
Gln	Asn
Glu	Asp
Gly	Pro
His	Asn, Gln
Ile	Leu, Val
Leu	Ile, Val
Lys	Arg, Gln, Glu
Met	Leu, Ile,
Phe	Met, Leu, Tyr
Ser	Thr
Thr	Ser
Trp	Tyr
Tyr	Trp, Phe
Val	Ile, Leu

La sustitución conservadora de aminoácidos también incluye agrupamientos basados en las cadenas laterales. Por ejemplo, un grupo de aminoácidos que tienen cadenas laterales alifáticas es glicina, alanina, valina, leucina e isoleucina; un grupo de aminoácidos que tienen cadenas laterales alifáticas-hidroxilo es serina y treonina; un grupo de aminoácidos que tienen cadenas laterales que contienen amida es asparagina y glutamina; un grupo de aminoácidos que tienen cadenas laterales aromáticas es fenilalanina, tirosina y triptófano; un grupo de aminoácidos que tienen cadenas laterales básicas es lisina, arginina e histidina; y un grupo de aminoácidos que tienen cadenas laterales que contienen azufre es cisteína y metionina. Por ejemplo, es razonable esperar que el reemplazo de una leucina por una isoleucina o valina, un aspartato por un glutamato, una treonina por una serina o un reemplazo similar de un aminoácido con un aminoácido estructuralmente relacionado no tendrá un efecto importante sobre las propiedades del polipéptido variante resultante. Se muestran sustituciones conservadoras en la Tabla 11 a continuación bajo el encabezamiento de sustituciones preferidas. Se muestran sustituciones más preferidas bajo el encabezamiento de sustituciones de ejemplo. Las sustituciones de aminoácidos, en general, se logran mediante la selección de sustituciones que no difieren significativamente en su efecto de mantener (a) la estructura de la cadena principal peptídica en el área de la sustitución, (b) la carga o hidrofobia de la molécula en el sitio diana o (c) el

grueso de la cadena lateral. Después de que se introducen las sustituciones, las variantes se seleccionan por su actividad biológica.

TABLA 11

SUSTITUCIONES DE AMINOÁCIDOS DE EJEMPLO Y PREFERIDAS		
Resto original	Sustituciones de ejemplo	Sustituciones preferidas
Ala	Val, Leu, Ile	Val
Arg	Lys, Gln, Asn	Lys
Asn	Gln, His, Lys, Arg	Gln
Asp	Glu	Glu
Cys	Ser	Ser
Gln	Asn, His, Lys,	Asn
Glu	Asp, Lys	Asp
Gly	Pro	Pro
His	Asn, Gln, Lys, Arg	Arg
Ile	Leu, Val, Met, Ala, Phe, Norleu	Leu
Leu	Norleu, Ile, Val, Met, Ala, Phe	Ile
Lys	Arg, Gln, Asn	Arg
Met	Leu, Ile, Phe	Leu
Phe	Leu, Val, Ile, Ala	Leu
Pro	Gly	Gly
Ser	Thr	Thr
Thr	Ser	Ser
Trp	Tyr	Tyr
Tyr	Trp, Phe, Thr, Ser	Phe
Val	Ile, Leu, Met, Phe, Ala, Norleu	Leu

Como alternativa, pueden agruparse aminoácidos similares para hacer sustituciones conservadoras en tres categorías basadas en la identidad de las cadenas laterales. El primer grupo incluye ácido glutámico, ácido aspártico, arginina, lisina, histidina, todos los cuales tienen cadenas laterales cargadas; el segundo grupo incluye glicina, serina, treonina, cisteína, tirosina, glutamina, asparagina; y el tercer grupo incluye leucina, isoleucina, valina, alanina, prolina, fenilalanina, triptófano, metionina, como se describe en Zubay, G., *Biochemistry*, tercera edición, Wm.C. Editores Brown (1993).

La divulgación incluye adicionalmente una construcción quimérica que comprende un polinucleótido sintético, que está unido operativamente a una secuencia reguladora. La secuencia reguladora comprende convenientemente secuencias de control de la transcripción y/o de la traducción, que serán compatibles para la expresión en el organismo de interés o en células de ese organismo. Normalmente, las secuencias de control reguladoras de la transcripción y de la traducción incluyen, pero no se limitan a, una secuencia promotora, una región 5' no codificante, una región reguladora en *cis* tal como un sitio de unión funcional para una proteína reguladora de la transcripción o una proteína reguladora de la traducción, un marco abierto de lectura aguas arriba, secuencias de unión ribosómica, un sitio de inicio de la transcripción, un sitio de inicio de la traducción y/o una secuencia de nucleótidos que codifica una secuencia líder, una codón de terminación, un sitio de parada de la traducción y una región 3' no traducida. Los promotores constitutivos o inducibles como se conocen en la técnica se incluyen en la invención. Los promotores pueden ser promotores de origen natural o promotores híbridos que combinan elementos de más de un promotor. Las secuencias promotoras incluidas en la presente invención pueden ser nativas para el organismo de interés o pueden derivar de una fuente alternativa, en las que la región es funcional en el organismo elegido. La elección del promotor diferirá dependiendo del hospedador o célula o tipo de tejido previstos. Por ejemplo, los promotores que podrían usarse para la expresión en mamíferos incluyen el promotor de metalotioneína, que puede inducirse en respuesta a metales pesados tales como cadmio, el promotor de β -actina, así como promotores víricos tales como el promotor de antígeno T grande SV40, el promotor temprano inmediato (IE) de citomegalovirus humano (CMV), el promotor LTR del virus del sarcoma de Rous, el promotor LTR del virus del tumor mamario de ratón, el promotor tardío principal de adenovirus (Ad MLP), el promotor del virus herpes simple y un promotor de VPH, en particular la región reguladora aguas arriba (URR, por sus siglas en inglés) del VPH, entre otros. Todos estos promotores están bien descritos y fácilmente disponibles en la técnica.

También pueden usarse elementos potenciadores en el presente documento para aumentar los niveles de expresión de las construcciones de mamíferos. Los ejemplos incluyen el potenciador del gen temprano de SV40, como se describe por ejemplo en Dijkema et al. (1985, *EMBO J.* 4:761), el potenciador/promotor derivado de la repetición terminal larga (LTR, por sus siglas en inglés) del Virus del Sarcoma de Rous, como se describe por ejemplo en Gorman et al., (1982, *Proc. Natl. Acad. Sci. EE.UU.* 79: 6777) y elementos derivados del CMV humano, como se describe por ejemplo en Boshart et al. (1985, *Cell* 41:521), tales como elementos incluidos en el intrón la secuencia de intrón A de CMV.

La construcción quimérica puede comprender también una secuencia no traducida 3'. La secuencia no traducida 3' se refiere a aquella porción de un gen que comprende un segmento de ADN que contiene una señal de poliadenilación y cualquier otra señal reguladora capaz de efectuar el procesamiento de ARNm o la expresión génica. La señal de poliadenilación se caracteriza por efectuar la adición de segmentos de ácido poliadenílico al extremo 3' del precursor de ARNm. Las señales de poliadenilación se reconocen habitualmente por la presencia de homología con la forma canónica 5' AATAAA-3' aunque las variaciones son frecuentes. Secuencia de ADN reguladora no traducida 3' preferentemente incluye de aproximadamente 50 a 1000 nucleótidos y puede contener secuencias de terminación de la transcripción y de la traducción además de a una señal de poliadenilación y cualquier otra señal reguladora capaz de efectuar el procesamiento de ARNm o la expresión génica.

En algunas realizaciones, la construcción quimérica contiene adicionalmente un gen marcador seleccionable para permitir la selección de células que contienen la construcción. Los genes de selección son bien conocidos en la técnica y serán compatibles para la expresión en la célula de interés.

Se entenderá, sin embargo, que la expresión de polinucleótidos que codifican proteínas en sistemas heterólogos ahora es bien conocida y la presente invención no se refiere a ni depende de ningún vector, secuencia de control de la transcripción o técnica para la expresión de los polinucleótidos particulares. Más bien, pueden introducirse polinucleótidos sintéticos preparados de acuerdo con los métodos que se exponen en el presente documento en un mamífero de cualquier forma adecuada en forma de cualquier construcción o vector adecuados y los polinucleótidos sintéticos pueden expresarse con elementos reguladores de la transcripción conocidos de cualquier forma convencional.

Además, pueden construirse construcciones quiméricas que incluyen secuencias que codifican adyuvantes. Son particularmente adecuados los mutantes desintoxicados de toxinas de ribosilación de ADP bacterianas, por ejemplo, la toxina diftérica, la toxina pertúsica (TP), la toxina colética (TC), las toxinas termolábiles de *Escherichia coli* (TL1 y TL2), la endotoxina A de *Pseudomonas*, las toxinas de C2y C3 de *Clostridium botulinum*, así como las toxinas de *C. perfringens*, *C. spiriforme* y *C. difficile*. En algunas realizaciones, las construcciones quiméricas incluyen secuencias que codifican mutantes desintoxicados de toxinas termolábiles de *E. coli*, tales como los mutantes desintoxicados TL-K63 y TL-R72, que se describen en la Patente de los EE.UU. N.º 6.818.222. En algunas realizaciones, el adyuvante es un elemento de desestabilización de proteínas, lo que aumenta el procesamiento y la presentación del polipéptido que corresponde a al menos una porción del antígeno diana a través de la vía del CMH de clase I, lo que conduce a la inmunidad mediada por células potenciada frente al polipéptido. Los elementos desestabilizadores de proteínas ilustrativos incluyen señales de degradación de proteínas intracelulares o degrones que pueden seleccionarse sin limitación entre un aminoácido desestabilizador en el extremo amino terminal de un polipéptido de interés, una región PEST o una ubiquitina. Por ejemplo, la secuencia codificante para el polipéptido puede modificarse para incluir un aminoácido desestabilizador en su extremo amino terminal de manera que la proteína modificada de este modo está sujeta a la vía de la regla del extremo N como se desvela, por ejemplo, por Bachmair et al. en la Patente de los EE.UU. N.º de serie 5.093.242 y por Varshavsky et al. en la Patente de los EE.UU. N.º de serie 5.122.463. En algunas realizaciones, el aminoácido desestabilizador se selecciona entre isoleucina y ácido glutámico, especialmente entre histidina, tirosina y glutamina y, más especialmente, entre ácido aspártico, asparagina, fenilalanina, leucina, triptófano y lisina. En ciertas realizaciones, el aminoácido desestabilizador es arginina. En algunas proteínas, el extremo amino terminal está oculto como resultado de la conformación de la proteína (es decir, su estructura terciaria o cuaternaria). En estos casos, puede ser necesario una alteración más extensa del extremo amino para hacer que la proteína se someta a la vía de la regla del extremo N. Por ejemplo, cuando la adición simple o el reemplazo del resto amino terminal solo es insuficiente debido a un extremo amino inaccesible, pueden añadirse varios aminoácidos (incluyendo lisina, el sitio de la ubiquitina que se une a proteínas sustrato) al extremo amino original para aumentar la accesibilidad y/o movilidad segmentaria del extremo amino modificado por ingeniería genética. En algunas realizaciones, una secuencia de ácido nucleico que codifica la región amino terminal del polipéptido puede modificarse para introducir un resto de lisina en un contexto apropiado. Esto puede lograrse más convenientemente mediante el empleo de construcciones de ADN que codifican "segmentos desestabilizadores universales". Un segmento desestabilizador universal, comprende una construcción de ácido nucleico que codifica una estructura de polipéptido, preferentemente segmentariamente móvil, que contiene uno o más restos de lisina, estando los codones para los restos de lisina posicionados dentro de la construcción de manera que cuando se inserta la construcción en la secuencia codificante del polinucleótido sintético que codifica proteína, los restos de lisina están suficientemente próximos espacialmente al extremo amino terminal de la proteína codificada para servir como el segundo factor determinante de la señal de degradación amino terminal completa. La inserción de dichas construcciones en la porción 5' de un polinucleótido sintético que codifica polipéptido proporcionaría el polipéptido codificado con un resto (o restos) de lisina en un contexto apropiado para la

desestabilización. En otras realizaciones, el polipéptido se modifica para contener una región PEST, que es rica en un aminoácido seleccionado entre prolina, ácido glutámico, serina y treonina, región que está opcionalmente flanqueada por aminoácidos que comprenden cadenas laterales electropositivas. En este sentido, se sabe que las secuencias de aminoácidos de las proteínas con semividas intracelulares inferiores aproximadamente 2 horas contienen una o más regiones ricas en prolina (P), ácido glutámico (E), serina (S) y treonina (T) como se muestra por ejemplo por Rogers et al. (1986, *Science* 234 (4774): 364-368). En otras realizaciones más, el polipéptido se conjuga con una ubiquitina o un fragmento biológicamente activo de la misma, para producir un polipéptido modificado cuya velocidad de degradación proteolítica intracelular aumenta, se potencia o se eleva de otra forma con respecto al polipéptido sin modificar.

Pueden coexpresarse uno o más polipéptidos adyuvantes con un polipéptido 'antigénico' que corresponde al menos una porción del antígeno diana. En ciertas realizaciones, pueden coexpresarse adyuvantes y polipéptidos antigénicos en forma de una proteína de fusión que comprende uno o más polipéptidos adyuvantes y uno o más polipéptidos antigénicos. Como alternativa, pueden coexpresarse adyuvantes y polipéptidos antigénicos como proteínas separadas.

Además, pueden construirse construcciones quiméricas que incluyen secuencias génicas de codificación de antígenos quiméricos, que codifican, por ejemplo, múltiples antígenos/epítomos de interés, por ejemplo, derivados de un único antígeno diana o de más de uno. En ciertas realizaciones, pueden construirse casetes multicistrónicos (por ejemplo, casetes bicistrónicos) que permiten la expresión de múltiples adyuvantes y/o polipéptidos antigénicos a partir de un único ARNm usando, por ejemplo, el ECMV IRES o similares. En otras realizaciones, pueden codificarse polipéptidos adyuvantes y/o antigénicos en secuencias de codificación separadas que están conectados operativamente a elementos reguladores de la transcripción independientes.

En algunas realizaciones, las construcciones quiméricas están en forma de vectores de expresión que se seleccionan convenientemente entre vectores extracromosómicos autorreplicantes (por ejemplo, plásmidos) y vectores que se integran en un genoma hospedador. En ejemplos ilustrativos de este tipo, los vectores de expresión son vectores víricos, tales como virus de simio 40 (VS40) o virus del papiloma bovino (VPB), que tienen la capacidad de replicarse como elementos extracromosómicos (*Eukaryotic Viral Vectors*, Cold Spring Harbor Laboratory, Gluzman ed., 1982; Sarver et al., 1981, *Mol. Cell Biol.* 1: 486). Los vectores víricos incluyen retrovirus (lentivirus), virus adenoasociado (véanse, por ejemplo, Okada, 1996, *Gene Ther.* 3: 957-964; Muzyczka, 1994, *J. Clin. Invest.* 94: 1351; las Patentes de los EE.UU. N.º 6.156.303; 6.143.548, 5.952.221, que describen vectores de VAA; véanse también las Patentes de los EE.UU. N.º 6.004.799; 5.833.993), adenovirus (véanse, por ejemplo, las Patentes de los EE.UU. N.º 6.140.087; 6.136.594; 6.133.028; 6.120.764), reovirus, herpesvirus, rotavirus etc., modificados para introducir y dirigir la expresión de un polinucleótido o transgén en las células. Los vectores retrovíricos pueden incluir aquellos basados en el virus de la leucemia murina (véase, por ejemplo, la Patente de los EE.UU. N.º 6.132.731), el virus de la leucemia de simio gibbon (véase, por ejemplo, la Patente de los EE.UU. N.º 6.033.905), el virus de la inmunodeficiencia de simio, el virus de la inmunodeficiencia humana (véase, por ejemplo, la Patente de los EE.UU. N.º 5.985.641) y combinaciones de los mismos.

Los vectores también incluyen aquellos que entregan genes eficientemente a células de animales *in vivo* (por ejemplo, células madre) (véase, por ejemplo, la Patente de los EE.UU. N.º 5.821.235 y 5.786.340; Croyle et al., 1998, *Gene Ther.* 5: 645; Croyle et al., 1998, *Pharm. Res.* 15: 1348; Croyle et al., 1998, *Hum. Gene Ther.* 9: 561; Foreman et al., 1998, *Hum. Gene Ther.* 9: 1313; Wirtz et al., 1999, *Gut* 44: 800). Los vectores víricos adenovíricos y adenoasociados adecuados para la entrega *in vivo* se describen, por ejemplo, en las Patentes de los EE.UU. N.º 5.700.470, 5.731.172 y 5.604.090. Los vectores adicionales adecuados para la entrega *in vivo* incluyen vectores del virus del herpes simple (véase, por ejemplo, la Patente de los EE.UU. N.º 5.501.979), vectores retrovíricos (véase, por ejemplo, las Patentes de los EE.UU. N.º 5.624.820, 5.693.508 y 5.674.703; y el documento WO92/05266 y el documento WO92/14829), vectores del virus del papiloma bovino (VPB) (véase, por ejemplo, la Patente de los EE.UU. N.º 5.719.054), vectores basados en el CMV (véase, por ejemplo, la Patente de los EE.UU. N.º 5.561.063) y vectores de parvovirus, rotavirus y virus de Norwalk. Los vectores lentivíricos son útiles para infectar tanto células en división como células que no están en división (véase, por ejemplo, la Patente de los EE.UU. N.º 6.013.516).

Los vectores víricos adicionales que encontrarán un uso para la entrega de las moléculas de ácido nucleico que codifican los antígenos de interés incluyen los derivados de la familia de virus de la viruela, incluyendo el virus variolovacunal y la viruela aviar. A modo de ejemplo, pueden construirse recombinantes de virus variolovacunal que expresan las construcciones quiméricas como se indica a continuación. La secuencia codificante de antígeno se inserta primero en un vector apropiado de manera que esté adyacente a un promotor de virus variolovacunal y secuencias de ADN de virus variolovacunal flanqueantes, tales como la secuencia que codifica la timidina cinasa (TK). Después, este vector se usa para transfectar células que se infectan simultáneamente con virus variolovacunal. La recombinación homóloga sirve para insertar el promotor de virus variolovacunal más el gen que codifica las secuencias codificantes de interés en el genoma vírico. El recombinante de TK resultante puede seleccionarse cultivando las células en presencia de 5-bromodesoxiuridina y recogiendo placas víricas resistentes a la misma.

Como alternativa, también pueden usarse avipoxvirus, tales como los virus de la viruela aviar y de la viruela del canario, para entregar los genes. Se sabe que los virus avipox recombinantes, que expresan inmunógenos de patógenos de mamíferos, confieren inmunidad protectora cuando se administran a especies no aviares. El uso de un vector de avipox es particularmente deseable en seres humanos y otras especies de mamíferos puesto que los miembros del género avipox solo pueden replicarse productivamente en especies aviares susceptibles y, por tanto, no son infecciosos en células de mamíferos. Se conocen en la técnica métodos para producir avipoxvirus recombinantes y emplean recombinación genética, como se ha descrito anteriormente con respecto a la producción de virus variolovacunales. Véanse, por ejemplo, el documento WO 91/12882; el documento WO 89/03429; y el documento WO 92/03545.

También puede usarse para la entrega génica vectores conjugados moleculares, tales como los vectores quiméricos de adenovirus descritos en Michael et al., *J. Biol. Chem.* (1993) 268: 6866-6869 y Wagner et al., *Proc. Natl. Acad. Sci. USA* (1992) 89: 6099-6103.

También encontrarán un uso como vectores víricos para la entrega de las construcciones quiméricas miembros del género Alphavirus, tales como, pero no limitados a, vectores derivados del virus Sindbis (SIN), el virus Semliki Forest (SFV) y el virus de la encefalitis equina venezolana (VEE). Para una descripción de vectores derivados del virus Sindbis útiles para la puesta en práctica de los métodos de la invención, véanse, Dubensky et al. (1996, *J. Virol* 70: 508-519; y las Publicaciones Internacionales N.º WO 95/07995, WO 96/17072); así como, Dubensky, Jr., T. W., et al., la Patente de los EE.UU. N.º 5.843.723 y Dubensky, Jr., T. W., Patente de los EE.UU. N.º 5.789.245. Son vectores ejemplares de este tipo vectores de alfavirus quiméricos que comprenden secuencias derivadas del virus Sindbis y del virus de la encefalitis equina venezolana. Véanse, por ejemplo, Perri et al. (2003, *J. Virol.* 77: 10394-10403) y las Publicaciones Internacionales N.º WO 02/099035, WO 02/080982, WO 01/81609 y WO 00/61772.

En otras realizaciones ilustrativas, se emplean vectores lentivíricos para entregar una construcción quimérica en células o tejidos seleccionados. Normalmente, estos vectores comprenden un LTR lentivírico 5', un sitio de unión a ARNt, una señal de empaquetamiento, un promotor unido operativamente a uno o más genes de interés, un origen de síntesis de la segunda cadena de ADN y un LTR lentivírico 3', en los que el vector lentivírico contiene un elemento de transporte nuclear. El elemento de transporte nuclear puede estar situado ya sea aguas arriba (5') o aguas abajo (3') con respecto a una secuencia codificante de interés (por ejemplo, un casete de expresión sintético Gag o Env). Puede usarse una amplia diversidad de lentivirus dentro del contexto de la presente invención, incluyendo, por ejemplo, lentivirus seleccionados entre el grupo que consiste en VIH, VIH-1, VIH-2, FIV, BIV, VAIE, MW, CAEV y SIV. Se describen ejemplos ilustrativos de vectores lentivíricos en las Publicaciones PCT N.º WO 00/66759, WO 00/00600, WO 99/24465, WO 98/51810, WO 99/51754, WO 99/31251, WO 99/30742 y WO 99/15641. Deseablemente, se usa un lentivirus SIN de tercera generación. Los proveedores comerciales de lentivirus SIN de tercera generación (autoinactivación) incluyen Invitrogen (ViraPower Lentiviral Expression System). Se proporcionan métodos detallados para la construcción, la transfección, la recolección y el uso de vectores lentivíricos, por ejemplo, en el manual técnico de Invitrogen "ViraPower Lentiviral Expression System version B 050102 25-0501", disponible en http://www.invitrogen.com/Content/Tech-Online/molecular_biology/manuals_pps/virapower_lentiviral_system_man.pdf. Los vectores lentivíricos han surgido como un método eficiente para la transferencia de genes. Las mejoras en las características de bioseguridad han hecho estos vectores adecuados para su uso a nivel de bioseguridad 2 (BL2). Una serie de características de seguridad se incorporan en vectores SIN de tercera generación (autoinactivación). La delección de la región 3' LTR U3 vírica da como resultado un provirus que es incapaz de transcribir un ARN vírico de longitud completa. Además, se proporciona una serie de genes esenciales en trans, produciendo una reserva vírica que es capaz de una sola ronda de infección e integración. Los vectores lentivíricos tienen varias ventajas, entre ellas: 1) la pseudotipificación del vector usando proteínas de la envoltura anfotrópicas les permite infectar virtualmente cualquier tipo celular; 2) se ha demostrado la entrega de genes a células quiescentes, post mitóticas, diferenciadas, incluyendo neuronas; 3) su baja toxicidad celular es única entre los sistemas de entrega transgénicos; 4) la integración vírica en el genoma permite la expresión del transgén a largo plazo; 5) su capacidad de empaquetamiento (6-14 kb) es mucho mayor que la de otros vectores retrovíricos o víricos adenoasociados. En una reciente demostración de las capacidades de este sistema, se usaron vectores lentivíricos que expresaban GFP para infectar células madre murinas dando como resultado progenie viva, transmisión de la estirpe germinal y expresión específica de promotor y de tejido del indicador (Ailles, L. E. y Naldini, L., *HIV-1-Derived Lentiviral Vectors*. En: Trono, D. (Ed.), *Lentiviral Vectors*, Springer-Verlag, Berlín, Heidelberg, Nueva York, 2002, págs. 31-52). Un ejemplo de los vectores de la generación actual se esboza en la FIG. 2 de una revisión de Lois et al. (2002, *Science*, 295 868-872).

La construcción quimérica también puede entregarse sin un vector. Por ejemplo, la construcción quimérica puede empaquetarse como ADN o ARN en liposomas antes de la entrega al sujeto o a células derivadas del mismo. La encapsulación en lípidos generalmente se consigue usando liposomas que son capaces de unirse de forma estable o atrapar y retener ácido nucleico. La relación de ADN condensado con preparación lipídica puede variar pero en general será de aproximadamente 1:1 (mg de ADN:micromoles de lípido) o más de lípido. Para una revisión del uso de liposomas como vehículos para la entrega de ácidos nucleicos, véanse, Hug y Sleight, (1991, *Biochim. Biophys. Acta*. 1097: 1-17) y Straubinger et al., en *Methods of Enzymology* (1983), vol. 101, págs. 512-527.

Las preparaciones de liposomas para su uso en la presente invención incluyen preparaciones catiónicas (cargadas positivamente), aniónicas (cargadas negativamente) y neutras, prefiriéndose en particular los liposomas catiónicos. Se ha demostrado que los liposomas catiónicos median en la administración intracelular de ADN plasmídico; (Felgner et al., 1987, *Proc. Natl. Acad. Sci. USA* 84: 7413-7416); ARNm (Malone et al., 1989, *Proc. Natl. Acad. Sci. USA* 86: 6077-6081); y factores de transcripción purificados (Debs et al., 1990, *J. Biol. Chem.* 265: 10189-10192), en forma funcional.

Los liposomas catiónicos están fácilmente disponibles. Por ejemplo, existen liposomas de N[1-2,3-dioleiloxi]propil]-N,N,N-trietilamonio (DOTMA) disponibles con la marca comercial Lipofectina, de GIBCO BRL, Grand Island, N.Y. (Véase, también, Felgner et al., 1987, *Proc. Natl. Acad. Sci. USA* 84: 7413-7416). Otros lípidos disponibles en el mercado incluyen (DDAB/DOPE) y DOTAP/DOPE (Boehringer). Pueden prepararse liposomas catiónicos alternativos a partir de materiales fácilmente disponibles usando técnicas bien conocidas en la técnica. Véanse, por ejemplo, Szoka et al., 1978, *Proc. Natl. Acad. Sci. USA* 75: 4194-4198; Publicación PCT N.º WO 90/11092 para una descripción de la síntesis de liposomas de DOTAP (1,2-bis(oleoiloxi)-3-(trimetilamonio)propano).

De forma similar, existen liposomas aniónicos y neutros fácilmente disponibles, tales como, de Avanti Polar Lipids (Birmingham, Ala) o pueden prepararse fácilmente usando materiales fácilmente disponibles. Dichos materiales incluyen fosfatidil colina, colesterol, fosfatidil etanolamina, dioleoilfosfatidil colina (DOPC), dioleoilfosfatidil glicerol (DOPG), dioleoilfosfatidil etanolamina (DOPE), entre otros. Estos materiales también pueden mezclarse con los materiales de partida DOTMA y DOTAP en proporciones adecuadas. Los métodos para fabricar liposomas usando estos materiales son bien conocidos en la técnica.

Los liposomas pueden comprender vesículas multilamelares (VML), vesículas unilaminares pequeñas (VUP) o vesículas unilamelares grandes (VUG). Los diversos complejos liposoma-ácido nucleico se preparan usando métodos conocidos en la técnica. Véanse, por ejemplo, Straubinger et al., en *METHODS OF IMMUNOLOGY* (1983), Vol. 101, págs. 512-527; Szoka et al., 1978, *Proc. Natl. Acad. Sci. USA* 75: 4194-4198; Papahadjopoulos et al., 1975, *Biochim. Biophys. Acta* 394: 483; Wilson et al., 1979, *Cell* 17:77; Deamer y Bangham, 1976, *Biochim. Biophys. Acta* 443: 629; Ostro et al., 1977, *Biochem. Biophys. Res. Commun.* 76: 836; Fraley et al., 1979, *Proc. Natl. Acad. Sci. USA* 76: 3348; Enoch y Strittmatter, 1979, *Proc. Natl. Acad. Sci. USA* 76: 145; Fraley et al., 1980, *J. Biol. Chem.* 255: 10431; Szoka y Papahadjopoulos, 1978, *Proc. Natl. Acad. Sci. USA* 75: 145; y Schaefer-Ridder et al., 1982, *Science* 215: 166.

La construcción quimérica también puede entregarse en composiciones lipídicas de cocleato similares a las descritas por Papahadjopoulos et al., 1975, *Biochem. Biophys. Acta*. 394: 483-491. Véanse, también, las Patentes de los EE.UU. N.º 4.663.161 y 4.871.488.

La construcción quimérica también puede encapsularse, adsorberse en o asociarse a vehículos en forma de partículas. Dichos vehículos presentan múltiples copias de una construcción quimérica seleccionada al sistema inmunitario. Las partículas pueden ser absorbidas por las células presentadoras de antígenos profesionales tales como macrófagos y células dendríticas, y/o pueden potenciar la presentación de antígenos a través de otros mecanismos tales como la estimulación de la liberación de citocinas. Los ejemplos de vehículos en forma de partículas incluyen los derivados de polímeros de metacrilato de polimetilo, así como micropartículas derivadas de poli(lactidas) y poli(lactida-co-glicolidas), conocidas como PLG. Véanse, por ejemplo, Jeffery et al., 1993, *Pharm. Res.* 10: 362-368; McGee J. P., et al., 1997, *J. Microencapsul.* 14 (2): 197-210; O'Hagan D. T., et al., 1993, *Vaccine* 11 (2): 149-54.

Además, pueden usarse otros sistemas en forma de partículas y polímeros para la entrega *in vivo* de la construcción quimérica. Por ejemplo, polímeros tales como polilisina, poliarginina, poliornitina, espermina, espermidina, así como conjugados de estas moléculas, son útiles para transferir un ácido nucleico de interés. De forma similar, la transfección mediada por DEAE dextrano, la precipitación con fosfato de calcio o la precipitación usando otras sales inorgánicas insolubles, tales como fosfato de estroncio, silicatos de aluminio incluyendo bentonita y caolín, óxido crómico, silicato de magnesio, talco y similares, encontrarán un uso con los presentes métodos. Véase, por ejemplo, Felgner, P. L., *Advanced Drug Delivery Reviews* (1990) 5: 163-187, para una revisión de sistemas de entrega útiles para la transferencia de genes. También puede usarse peptoides (Zuckerman, R. N., et al., Pat. N.º 5.831.005, expedida el 3 de noviembre de 1998) para la entrega de una construcción de la presente invención.

Adicionalmente, sistemas de entrega biolísticos que emplean vehículos en forma de partículas tales como oro y tungsteno, son especialmente útiles para la entrega de construcciones quiméricas. Las partículas se recubren con el casete o casetes de expresión sintética que se han de entregar y se aceleran a alta velocidad, generalmente a una atmósfera reducida, usando una descarga de pólvora de una "pistola de genes". Para una descripción de dichas técnicas y aparatos útiles para las mismas, véanse, por ejemplo, las Patentes de los EE.UU. N.º 4.945.050; 5.036.006; 5.100.792; 5.179.022; 5.371.015; y 5.478.744. En ejemplos ilustrativos, la aceleración de partículas accionada por gas puede conseguirse con dispositivos tales como los fabricados por PowderMed Pharmaceuticals PLC (Oxford, Reino Unido) y PowderMed Vaccines Inc. (Madison, Wis.), algunos ejemplos de los cuales se describen en las Patentes de los EE.UU. N.º 5.846.796; 6.010.478; 5.865.796; 5.584.807; y en la Patente EP N.º 0 500 799. Este enfoque ofrece un enfoque de entrega sin aguja en el que una formulación de polvo seco de

partículas microscópicas, tales como partículas de polinucleótido o polipéptido, se aceleran a alta velocidad dentro de un chorro de helio gaseoso generado por un dispositivo de mano, propulsando las partículas en un tejido diana de interés. Otros dispositivos y métodos que pueden ser útiles para la inyección sin aguja impulsada por gas de composiciones incluyen los proporcionados por Bioject, Inc. (Portland, Oreg.), algunos ejemplos de los cuales se describen en las Patentes de los EE.UU. N.º 4.790.824; 5.064.413; 5.312.335; 5.383.851; 5.399.163; 5.520.639 y 5.993.412.

Como alternativa, pueden usarse dispositivos basados en microcánulas y microagujas (tales como los que están siendo desarrollados por Becton Dickinson y otros) para administrar las construcciones quiméricas. Se describen dispositivos ilustrativos de este tipo en el documento EP 1 092 444 A1 y la Solicitud de los EE.UU. N.º de serie 606.909, presentada el 29 de junio de 2000. También puede usarse una cánula de acero convencional para la entrega intradérmica usando dispositivos y métodos como se describen en la Solicitud de los EE.UU. N.º de serie 417.671, presentada el 14 de octubre de 1999. Estos métodos y dispositivos incluyen la entrega de sustancias a través de una "micro-cánula" de calibre estrecho (aproximadamente 30 G) con profundidad de penetración limitada, definida por la longitud total de la cánula o la longitud total de la cánula que se expone más allá de una característica limitante de la profundidad. Está dentro del alcance de la presente invención que pueda conseguirse la administración dirigida de sustancias incluyendo construcciones quiméricas ya sea a través de una única microcánula o una matriz de microcánulas (o "microagujas"), por ejemplo 3-6 microagujas montadas en un dispositivo de inyección que puede incluir o estar unido a un depósito en el que está contenida la sustancia que se ha de administrar.

10. Composiciones

La divulgación también proporciona composiciones, en particular composiciones inmunomoduladoras, que comprenden una o más de las construcciones quiméricas que se describen en el presente documento. Las composiciones inmunomoduladoras pueden comprender una mezcla de construcciones quiméricas, que a su vez pueden entregarse, por ejemplo, usando los mismos o diferentes vectores o vehículos. Pueden administrarse antígenos individualmente o en combinación, en, por ejemplo, composiciones inmunomoduladoras profilácticas (es decir, para prevenir la infección o enfermedad) o terapéuticas (para tratar la infección o enfermedad). Las composiciones inmunomoduladoras pueden proporcionarse más de una vez (por ejemplo, una administración de "sensibilización" seguida de uno o más "refuerzos") para conseguir los efectos deseados. La misma composición puede administrarse en una o más sensibilizaciones y una o más etapas de refuerzo. Como alternativa, pueden usarse diferentes composiciones para la sensibilización y el refuerzo.

Las composiciones inmunomoduladoras incluirán en general uno o más "excipientes o vehículos farmacéuticamente aceptables", tales como agua, solución salina, glicerol, etanol, etc. Adicionalmente, pueden estar presentes en dichos vehículos sustancias auxiliares, tales como agentes humectantes o emulsionantes, sustancias tamponantes del pH y similares.

Las composiciones inmunomoduladoras, normalmente, además de los componentes mencionados anteriormente, comprenderán uno o más "vehículos farmacéuticamente aceptables". Estos incluyen cualquier vehículo que no induzca por sí mismo la producción de anticuerpos perjudiciales para el individuo que recibe la composición. Los vehículos adecuados normalmente son macromoléculas grandes, metabolizadas lentamente, tales como proteínas, polisacáridos, ácidos polilácticos, ácidos poliglicólicos, aminoácidos poliméricos, copolímeros de aminoácidos y agregados lipídicos (tales como gotitas de aceite o liposomas). Dichos vehículos son bien conocidos por los expertos habituales en la materia. Una composición puede contener también un diluyente, tal como agua, solución salina, glicerol, etc. Adicionalmente, pueden estar presentes una sustancia adyuvante, tal como un agente humectante o emulsionante, una sustancia tamponante del pH y similares. Hay disponible un análisis exhaustivo de componentes farmacéuticamente aceptables en Gennaro (2000) *Remington: The Science and Practice of Pharmacy*. 20ª ed, ISBN: 0683306472.

También pueden usarse sales farmacéuticamente compatibles en las composiciones, por ejemplo, sales minerales tales como clorhidratos, bromhidratos, fosfatos o sulfatos, así como sales de ácidos orgánicos tales como acetatos, propionatos, malonatos o benzoatos. Son sustratos proteicos especialmente útiles albúminas séricas, hemocianina de lapa californiana, moléculas de inmunoglobulina, tiroglobulina, ovoalbúmina, toxoide tetánico y otras proteínas bien conocidas por los expertos en la materia.

Las construcciones quiméricas también pueden adsorberse en, quedar atrapadas dentro o asociarse de otro modo a liposomas y vehículos en partículas tales como PLG.

Las construcciones quiméricas se formulan en composiciones para la entrega a un mamífero. Estas composiciones pueden ser ya sea profilácticas (para prevenir la infección) o terapéuticas (para tratar la enfermedad después de la infección). Las composiciones comprenderán una "cantidad terapéuticamente eficaz" del gen de interés de manera que pueda producirse una cantidad del antígeno *in vivo* de manera que se genere una respuesta inmunitaria en el individuo al que se administran. La cantidad exacta necesaria variará dependiendo del sujeto que se está tratando; de la edad y del estado general del sujeto que se trata; de la capacidad del sistema inmunitario del sujeto para

5 sintetizar anticuerpos; del grado de protección deseado; de la gravedad de la afección que se trata; del antígeno particular seleccionado y de su modo de administración, entre otros factores. Una cantidad eficaz apropiada puede ser determinada fácilmente por un experto en la materia. Por tanto, una "cantidad terapéuticamente eficaz" caerá en un intervalo relativamente amplio que puede determinarse mediante ensayos habituales.

10 Una vez formuladas, las composiciones pueden administrarse directamente al sujeto (por ejemplo, como se ha descrito anteriormente). La entrega directa de las composiciones que contienen construcción quimérica *in vivo* en general se realizará con o sin vectores, como se ha descrito anteriormente, por inyección usando una jeringa convencional, dispositivos sin aguja tales como Bioject™ o una pistola de genes, tal como el sistema de entrega de genes Accell™ (PowderMed Ltd, Oxford, Inglaterra) o un dispositivo de microagujas. Las construcciones pueden entregarse (por ejemplo, inyectarse) ya sea por vía subcutánea, por vía epidérmica, por vía intradérmica, por vía intramuscular, por vía intravenosa, por vía intramucosa (por ejemplo, por vía nasal, rectal y vaginal), por vía intraperitoneal o por vía oral. Se prefiere en particular la entrega de ácido nucleico en células de la epidermis ya que este modo de administración proporciona acceso a las células linfoides asociadas a la piel y proporciona una presencia transitoria de ácido nucleico (por ejemplo, ADN) en el receptor. Otros modos de administración incluyen la ingestión oral y la administración pulmonar, supositorios, inyección sin aguja, aplicaciones transcutáneas, tópicas y transdérmicas. El tratamiento de dosificación puede ser una posología de dosis única o una posología de dosis múltiples.

20 Con el fin de que la invención pueda comprenderse fácilmente y ponerse en efecto práctico, a continuación se describirán realizaciones preferidas particulares a modo de los siguientes ejemplos no limitantes.

Ejemplos

25 EJEMPLO 1

SISTEMA DE CONSTRUCCIONES SINTÉTICO PARA DETERMINAR LA PREFERENCIA DE RESPUESTA INMUNITARIA DE LOS CODONES EN MAMÍFEROS

30 Materiales y Métodos

Diseño/síntesis de cebadores y manipulación de secuencias

35 Se diseñaron oligonucleótidos para mutagénesis dirigida al sitio de acuerdo con las directrices incluidas en los manuales del kit de mutagénesis (kit de mutagénesis dirigida al sitio Quikchange II o kit de mutagénesis dirigida a múltiples sitios QuikChange; Stratagene, La Jolla CA). Estos cebadores se sintetizaron y se purificaron en PAGE por Sigma (anteriormente Proligo).

40 Se diseñaron a ojo oligonucleótidos para la síntesis del gen entero y se sintetizaron por Sigma (anteriormente Proligo). Los cebadores se suministraron como oligos desalados convencionales. No se realizó ninguna purificación adicional de los oligonucleótidos.

45 Se realizaron la manipulación y el análisis de la secuencia mediante el paquete de programas de Biomanager (ANGIS) y otros programas diversos en la web incluyendo BLAST en el NCBI (<http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi>), NEBcutter V2.0 de New England Biolabs (<http://tools.neb.com/NEBcutter2/index.php>), la Translate Tool en ExPASy (<http://au.expasy.org/tools/dna.html>) y el servidor SignalP 3.0 (<http://www.cbs.dtu.dk/services/SignalP/>).

Técnicas de clonación convencionales

50 Se realizaron digestiones de enzimas de restricción, tratamientos con fosfatasa alcalina y ligaciones de acuerdo con las instrucciones de los fabricantes de las enzimas (diversos fabricantes, incluyendo New England Biolabs, Roche y Fermentas).

55 Se realizaron la purificación de ADN a partir de geles de agarosa y la preparación de ADN minipreparativo usando kits comerciales (Qiagen, Bio-Rad, Macherey-Nagel).

60 Se realizaron la electroforesis en gel de agarosa, la extracción con fenol/cloroformo de proteínas contaminantes del ADN, la precipitación con etanol del ADN y otros procedimientos de biología molecular básicos usando protocolos convencionales, similares a los descritos en *Current Protocols in Molecular Biology* (ebook disponible a través de Wiley InterScience; editado por Ausubel *et al.*).

La secuenciación fue realizada por el Australian Genome Research Facility (Centro australiano para la investigación del genoma) (AGRF, Brisbane).

Síntesis del gen entero

Se usaron oligonucleótidos de ~35-50 monómeros (Sigma-Proligo) solapados para sintetizar secuencias de ADN más largas. Se incorporaron sitios de enzimas de restricción para facilitar la clonación. El método utilizado para sintetizar los fragmentos se basa en el que se proporciona en Smith *et al.* (2003). En primer lugar, los oligonucleótidos para la cadena superior o inferior se mezclaron y después se fosforilaron usando polinucleótido cinasa T4 (PNK; New England Biolabs). Las mezclas de oligonucleótidos después se purificaron a partir de la PNK mediante una extracción con fenol/cloroformo convencionales y precipitación con acetato de sodio/etanol (NaAc/EtOH). Después se mezclaron volúmenes iguales de las mezclas de oligonucleótidos para las cadenas superior e inferior y los oligonucleótidos se desnaturalizaron por calentamiento a 95 °C durante 2 minutos. Los oligonucleótidos se hibridaron mediante enfriamiento lento de la muestra a 55 °C y los oligonucleótidos hibridados se ligaron usando Taq ligasa (New England Biolabs). El fragmento resultante se purificó mediante extracción con fenol/CHCl₃ y precipitación con NaAc/EtOH.

Los extremos de los fragmentos se rellenaron y después los fragmentos se amplificaron, usando los cebadores directo e inverso más exteriores, con el kit de PCR Clontech Advantage HF 2 (Clontech) de acuerdo con las instrucciones del fabricante. Para rellenar los extremos se utilizó la siguiente PCR: 35 ciclos de una etapa de desnaturalización de 94 °C durante 15 s, una etapa de hibridación lenta en la que se redujo la temperatura a 55 °C durante 7 minutos y después se mantuvo a 55 °C durante 2 min y una etapa de elongación de 72 °C durante 6 minutos. Después se realizó una etapa de elongación final de 7 min a 72 °C. La segunda PCR para amplificar el fragmento implicó: una etapa de desnaturalización inicial a 94 °C durante 30 s, seguida de 25 ciclos de 94 °C durante 15 s, 55 °C durante 30 s y 68 °C durante 1 min y una etapa de elongación final de 68 °C durante 3 minutos.

Después los fragmentos se purificaron mediante electroforesis en gel, se digirieron y se ligaron en el vector correspondiente. Después de la transformación de *E. coli* con la mezcla de ligadura, se hicieron minipreparaciones para múltiples colonias y los insertos secuenciados. En ocasiones no fue posible aislar clones con la secuencia totalmente correcta. En esos casos los errores se fijaron por mutagénesis dirigida a uno o múltiples sitios.

Mutagénesis dirigida al sitio

La mutagénesis se realizó usando el kit de mutagénesis dirigida al sitio QuikChange II o el kit de mutagénesis dirigida a múltiples sitios QuikChange (Stratagene, La Jolla CA), con cebadores apropiados (Sigma) purificados por PAGE (electroforesis en gel de poliacrilamida), de acuerdo con las instrucciones del fabricante.

Preparación de construcciones

Los detalles de las construcciones utilizadas para generar la tabla de preferencias de codón se resumen en la TABLA 12. Todas las construcciones se fabricaron usando pCDNA3 de Invitrogen y se verificaron mediante secuenciación antes de su uso.

TABLA 12

SUMARIO DE LA CONSTRUCCIÓN DE E7 SECRETORA SERIES 1 Y 2				
Construcción	AA y codón	UC de la Sec sec	UC de E7	Proteína E7
Construcciones de control				
IgkC1	N/A	ts	ts	no onc
IgkC2	N/A	mc	cm	no onc
IgkC3	N/A	ts	ts	onc
IgkC4	N/A	mc	cm	onc
Construcción de E7 secretora serie 1				
IgkS1-1	Ala GCG	ts	ts con todo Ala gcg	no onc
IgkS1-2	Ala GCA	ts	ts con todo Ala gca	no onc
IgkS1-3	Ala GCT	ts	ts con todo Ala gct	no onc
IgkS1-4	Ala GCC	ts	ts con todo Ala gcc	no onc
IgkS1-5	Arg AGG	ts	ts con todo Arg agg	no onc
IgkS1-6	Arg AGA	ts	ts con todo Arg aga	no onc
IgkS1-7	Arg CGG	ts	ts con todo Arg cgg	no onc
IgkS1-8	Arg CGA	ts	ts con todo Arg cga	no onc
IgkS1-9	Arg CGT	ts	ts con todo Arg	no onc
			cgt	
IgkS1-10	Arg CGC	ts	ts con todo Arg cgc	no onc

Construcción de E7 secretora serie 1				
IgkS1-11	AsnAAT	ts	ts con todo Asn aat	no onc
IgkS1-12	Asn AAC	ts	ts con todo Asn aac	no onc
IgkS1-13	Asp GAT	ts con todo Asp gat	ts con todo Asp gat	no onc
IgkS1-14	Asp GAC	ts con todo Asp gac	ts con todo Asp gac	no onc
IgkS1-15	Cys TGT	ts	ts con todo Cys tgt	no onc
IgkS1-16	Cys TGC	ts	ts con todo Cys tgc	no onc
IgkS1-17	Glu GAG	ts con todo Glu gag	ts con todo Glu gag	no onc
IgkS1-18	Glu GAA	ts con todo Glu gaa	ts con todo Glu gaa	no onc
IgkS1-19	Gln CAG	ts	ts con todo Gln cag	no onc
IgkS1-20	Gln CAA	ts	ts con todo Gln caa	no onc
IgkS1-21	Gly GGG	ts con todo Gly ggg	ts con todo Gly ggg	no onc
IgkS1-22	Gly GGA	ts con todo Gly gga	ts con todo Gly gga	no onc
IgkS1-23	Gly GGT	ts con todo Gly ggt	ts con todo Gly ggt	no onc
IgkS1-24	Gly GGC	ts con todo Gly ggc	ts con todo Gly ggc	no onc
IgkS1-25	His CAT	ts	ts con todo His cat	no onc
IgkS1-26	His CAC	ts	ts con todo His cac	no onc
IgkS1-27	Ile ATA	ts	ts con todo Ile ata	no onc
IgkS1-28	Ile ATT	ts	ts con todo Ile att	no onc
IgkS1-29	Ile ATC	ts	ts con todo Ile atc	no onc
IgkS1-30	Lys AAG	ts	ts con todo Lys aag	no onc
IgkS1-31	Lys AAA	ts	ts con todo Lys aaa	no onc
IgkS1-32	Phe TTT	ts	ts con todo Phe ttt	no onc L15F, L22F
IgkS1-33	Phe TTC	ts	ts con todo Phe ttc	no onc L15F, L22F
IgkS1-34	Ser AGT	ts con todo Ser agt	ts con todo Ser agt	no onc
IgkS1-35	Ser AGC	ts con todo Ser age	ts con todo Ser agc	no onc
IgkS1-36	Ser TCG	ts con todo Ser tcg	ts con todo Ser teg	no onc
IgkS1-37	Ser TCA	ts con todo Ser tea	ts con todo Ser tca	no onc
IgkS1-38	Ser TCT	ts con todo Ser tct	ts con todo Ser tct	no onc
IgkS1-39	Ser TCC	ts	ts con todo Ser tee	no onc
IgkS1-40	Thr ACG	ts con todo Thr acg	ts con todo Thr acg	no onc
IgkS1-41	Thr ACA	ts con todo Thr aca	ts con todo Thr aca	no onc
IgkS1-42	Thr ACT	ts con todo Thr act	ts con todo Thr act	no onc
IgkS1-43	Thr ACC	ts con todo Thr acc	ts con todo Thr acc	no onc
IgkS1-44	Tyr TAT	ts	ts con todo Tyr tat	no onc
IgkS1-45	Tyr TAC	ts	ts con todo Tyr tac	no onc
IgkS1-46	Val GTG	ts con todo Val gtg	ts con todo Val gtg	no onc
IgkS1-47	Val GTA	ts con todo Val gta	ts con todo Val gta	no onc
IgkS1-48	Val GTT	ts con todo Val gtt	ts con todo Val gtt	no onc
IgkS1-49	Val GTC	ts con todo Val gtc	ts con todo Val gtc	no onc
IgkS1-50	Leu CTG	alterado con Leu ctg	alterado con Leu ctg	onc
IgkS1-51	Leu CTA	alterado con Leu cta	alterado con Leu cta	onc
IgkS1-52	Leu CTT	alterado con Leu ctt	alterado con Leu ctt	onc
IgkS1-53	Leu CTC	alterado con Leu ctc	alterado con Leu ctc	onc
IgkS1-54	Leu TTG	alterado con Leu ttg	alterado con Leu ttg	onc
IgkS1-55	Leu TTA	alterado con Leu tta	alterado con Leu tta	onc
IgkS1-56	Pro CCG	alterado con Pro ccg	alterado con Pro ccg	onc
IgkS1-57	Pro CCA	alterado con Pro cca	alterado con Pro cca	onc
IgkS1-58	ProCCT	alterado con Pro cct	alterado con Pro cct	onc

Construcción de E7 secretora serie 1				
IgkS1-59	ProCCC	alterado con Pro ccc	alterado con Pro ccc	onc
Construcción de E7 secretora serie 2				
IgkS2-1	Ala GCG	cm	cm	enlazadorA-onc
IgkS2-2	Ala GCA	cm	cm	enlazadorA-onc
IgkS2-3	Ala GCT	cm	cm	enlazadorA-onc
IgkS2-4	Ala GCC	cm	cm	enlazadorA-onc
IgkS2-5	Arg AGG	cm	cm	enlazadorR-onc
IgkS2-6	Arg AGA	cm	cm	enlazadorR-onc
IgkS2-7	Arg CGG	cm	cm	enlazadorR-onc
IgkS2-8	Arg CGA	cm	cm	enlazadorR-onc
IgkS2-9	Arg CGT	cm	cm	enlazadorR-onc
IgkS2-10	Arg CGC	cm	cm	enlazadorR-onc
IgkS2-11	Asn AAT	cm	cm	enlazadorN-onc
IgkS2-12	Asn AAC	cm	cm	enlazadorN-onc
IgkS2-13	Asp GAT	ts con todo Asp gat	ts con todo Asp gat	onc
IgkS2-14	Asp GAC	ts con todo Asp gac	ts con todo Asp gac	onc
IgkS2-15	Cys TGT	ts	ts con todo Cys tgt	onc
IgkS2-16	Cys TGC	ts	ts con todo Cys tgc	onc
IgkS2-17	Glu GAG	ts con todo Glu gag	ts con todo Glu gag	onc
IgkS2-18	Glu GAA	ts con todo Glu gaa	ts con todo Glu gaa	onc
IgkS2-19	Gln CAG	ts	ts con todo Gln cag	onc
IgkS2-20	Gln CAA	ts	ts con todo Gln caa	onc
IgkS2-21	Gly GGG	ts con todo Gly ggg	ts con todo Gly ggg	onc
IgkS2-22	Gly GGA	ts con todo Gly gga	ts con todo Gly gga	onc
IgkS2-23	Gly GGT	ts con todo Gly ggt	ts con todo Gly ggt	onc
IgkS2-24	Gly GGC	ts con todo Gly ggc	ts con todo Gly ggc	onc
IgkS2-25	His CAT	cm	cm	enlazadorH-onc
IgkS2-26	His CAC	cm	cm	enlazadorH-onc
IgkS2-27	Ile ATA	ts	ts con todo Ile ata	onc
IgkS2-28	Ile ATT	ts	ts con todo Ile att	onc
IgkS2-29	Ile ATC	ts	ts con todo Ile ate	onc
IgkS2-30	Lys AAG	cm	cm	enlazadorK-onc
IgkS2-31	Lys AAA	cm	cm	enlazadorK-onc
IgkS2-32	Phe TTT	cm	cm	enlazadorF-onc
IgkS2-33	Phe TTC	cm	cm	enlazadorF-onc
IgkS2-34	Ser AGT	ts con todo Ser agt	ts con todo Ser agt	onc
IgkS2-35	Ser AGC	ts con todo Ser agc	ts con todo Ser agc	onc
IgkS2-36	Ser TCG	ts con todo Ser tcg	ts con todo Ser tcg	onc
IgkS2-37	Ser TCA	ts con todo Ser tea	ts con todo Ser tea	onc
IgkS2-38	Ser TCT	ts con todo Ser tct	ts con todo Ser tct	onc
IgkS2-39	Ser TCC	ts	ts con todo Ser tcc	onc
IgkS2-40	Thr ACG	ts con todo Thr acg	ts con todo Thr acg	onc
IgkS2-41	Thr ACA	ts con todo Thr aca	ts con todo Thr aca	onc
IgkS2-42	Thr ACT	ts con todo Thr act	ts con todo Thr act	onc
IgkS2-43	Thr ACC	ts con todo Thr acc	ts con todo Thr acc	onc
IgkS2-44	Tyr TAT	cm	cm	enlazadorY-onc
IgkS2-45	Tyr TAC	cm	cm	enlazadorY-onc
IgkS2-46	Val GTG	ts con todo Val gtg	ts con todo Val gtg	onc
IgkS2-47	Val GTA	ts con todo Val gta	ts con todo Val gta	onc

Construcción de E7 secretora serie 2				
IgkS2-48	Val GTT	ts con todo Val gtt	ts con todo Val gtt	onc
IgkS2-49	Val GTC	ts con todo Val gtc	ts con todo Val gtc	onc
IgkS2-11b	Asn AAT	ts	ts con todo Asn aat	enlazadorN-no onc
IgkS2-12b	Asn AAC	ts	ts con todo Asn aac	enlazadorN-no onc
AA = aminoácido, UC = uso de consón, cm = consenso de mamífero, ts = tipo silvestre, onc = oncogénica, no onc = no oncogénica, Sec sec = secuencia secretora, N/A= no aplicable				

Construcciones de control

- 5 Las construcciones de control E7 se basaron en las de Liu *et al.* (2002). Se fabricaron construcciones de control E7 tanto oncogénicas (es decir, de tipo silvestre) como no oncogénicas con el uso de codón de tipo silvestre o de consenso de mamífero. E7 "no oncogénica" es E7 con mutaciones D21G, C24G, E26G, es decir, con mutaciones que se ha notificado que convierten E7 en no transformante (Edmonds y Vousden, 1989; Heck *et al.*, 1992).

- 10 La secuencia secretora derivó de ARN de IgK de *Mus musculus* para la cadena ligera del anticuerpo anti-HLA-DR (GenBank número de acceso D84070). Para algunas construcciones se modificó el uso del codón de esta secuencia.

Construcciones de control de uso de codón de tipo silvestre:

- 15 La construcción de E7 de uso de codón de tipos silvestre (ts) de Liu *et al.* se usó como molde en una PCR de mutagénesis dirigida al sitio para fabricar la construcción de E7 no oncogénica de uso de codón ts.

- 20 Las secuencias de E7 de uso de codón de tipo silvestre no oncogénica y oncogénica se amplificaron para incorporar un sitio 5' *Bam*HI y un sitio 3' *Eco*RI. Los fragmentos resultantes se clonaron en pCDNA3 con *Bam*HI y *Eco*RI cortados y se secuenciaron. El fragmento secretor se fabricó por síntesis del gen entero usando el uso de codón de tipo silvestre con sitios *Kpn*I y *Bam*HI flanqueantes. Los fragmentos secretores de Kozak después se ligaron en pCDNA3-tsE7 (no oncogénico o oncogénico) con *Kpn*ILBAMHI cortado para fabricar pCDNA3-Igk-nE7 y pCDNA3-Igk-E7 (nombrados IgkC1 y IgkC3 respectivamente; véase la TABLA 12). La identidad de las construcciones se confirmó por secuenciación.

Construcciones de control de uso de codón de consenso de mamífero (cm):

- 30 Como había errores en la construcción de E7 original de consenso de mamífero (cm) (L28F, Q70R y una delección de E35; Liu *et al.*, 2002) no se usó. Se sintetizó una construcción de control E7 no oncogénica de cm mediante síntesis de gen entero. Se fabricó una construcción de control E7 oncogénica cm (es decir, E7 de tipo silvestre) posteriormente a partir de la construcción de E7 no oncogénica cm mediante mutagénesis dirigida a un único sitio.

- 35 Se fabricaron construcciones oncogénicas y no oncogénicas cm secretoras mediante la amplificación de la secuencia de E7 cm con un cebador directo que introdujo un sitio *Bam*HI y un cebador inverso que incorporaba un sitio *Eco*RI. El fragmento de E7 resultante se clonó en los sitios respectivos en pCDNA3 y se secuenció. Una secuencia secretora cm flanqueada por los sitios *Kpn*I y *Bam*HI, 5' y 3' respectivamente, se sintetizó y se ligó en los sitios *Kpn*I y *Bam*HI de pCDNA3-cmE7 (oncogénico o no oncogénico) para fabricar pCDNA3-cmlgk-cmE7 y pCDNA3-cmlgk cmE7 (nombrados IgkC2 y IgkC4 respectivamente; véase la TABLA 12). La identidad de las construcciones se confirmó por secuenciación.

Construcciones de E7 no oncogénicas secretadas con uso de codón predominantemente de tipo silvestre, modificadas para codones individuales

- 45 Se fabricaron plásmidos que codificaban una forma no oncogénica de E7 para todos los codones, con la excepción de los codones de Pro y Leu, los codones de parada y los codones para aminoácidos no degenerados. Como Phe aparece solo una vez en la secuencia de E7, los codones para dos restos de Leu, L15 y L22, se mutaron a codones de Phe. Se usó una combinación de técnicas para fabricar estas construcciones. Cuando fueron necesarias pocas mutaciones se realizó una mutagénesis dirigida a uno o múltiples sitios de una construcción de control que codificaba E7 no oncogénica (los detalles de la construcción de control se han proporcionado anteriormente en "construcciones de control"). Cuando fueron necesarias modificaciones más extensas se empleó la síntesis de gen entero. Independientemente de los métodos utilizados, todas estas construcciones incluyen una secuencia que codifica E7 con idéntica secuencia aguas arriba y aguas abajo clonada en los sitios *Kpn*I y *Eco*RI de pCDNA3. Estas construcciones después se modificaron para incluir una secuencia secretora, como se describe a continuación.

- 55 En primer lugar, usando el método de síntesis de genes completos, se sintetizaron fragmentos de ADN que incluían una secuencia secretora flanqueada por sitios *Kpn*I y *Bam*HI. Para algunas construcciones el aminoácido de interés

apareció en la secuencia secretora de modo que se fabricaron fragmentos de secuencias secretoras modificadas individuales. Para construcciones para aminoácidos que no aparecieron en la secuencia secretora, se usó la secuencia secretora de tipo silvestre. Estos fragmentos se digirieron con *KpnI* y *BamHI*. Después, usando la construcción nE7 pertinente como molde y un protocolo de PCR convencional, se introdujo un sitio *BamHI* en el extremo 5' de la secuencia E7. El sitio *EcoRI* 3' se conservó. Los fragmentos E7 resultantes se cortaron con *BamHI* y *EcoRI*, se purificaron y se ligaron en pCDNA3. Después de la secuenciación, los plásmidos se cortaron con *KpnI* y *BamHI* y se ligaron con las secuencias secretoras *KpnI/BamHI* pertinentes. Las secuencias de las construcciones después se confirmaron. Se fabricaron las construcciones IgkS1-1 a IgkS1-49 de esta manera (véanse la TABLA 12 y las Figuras 1 a 11, 13 y 15 a 17 para comparaciones de secuencias).

Construcciones de E7 secretadas con codones de Pro o Leu individuales modificados

Se diseñaron secuencias de ADN de E7 en las que se modificaron codones de Pro o Leu individualmente. El resto del uso de codón para estos ADN de E7 fue el mismo para todas las construcciones de Pro y Leu, pero difirió del uso de codón de tipo silvestre o de consenso de mamífero. [Nótese que este uso de codón se basó en los datos preliminares de la inmunización de ratones con las construcciones GFP].

Los fragmentos de ADN de E7 Pro/Leu, flanqueados por los sitios *HindIII* y *BamHI*, se fabricaron mediante síntesis de gen entero y se clonaron en los sitios *HindIII* y *BamHI* de pCDNA3. Usando estas construcciones como moldes, se incorporó un sitio *KpnI* aguas arriba y un sitio *EcoRI* aguas abajo, de las secuencias de E7 Pro/Leu mediante métodos de PCR convencionales. Los fragmentos resultantes se cortaron con *KpnI* y *EcoRI* y se clonaron en pCDNA3. Estas construcciones después se usaron para fabricar las construcciones de E7 secretadas con modificaciones de codón de Pro o Leu.

En primer lugar, usando el método de síntesis de gen entero, se sintetizaron fragmentos de ADN que incluían una secuencia secretora flanqueada por sitios *KpnI* y *BamHI*. Como Pro y Leu aparecen en la secuencia secretora, se fabricaron fragmentos de secuencia secretora modificados individualmente para las diferentes construcciones. Estos fragmentos se digirieron con *KpnI* y *BamHI*. Después, usando la construcción de E7 de Pro o Leu pertinente como un molde y un protocolo de PCR convencional, se introdujo un sitio *BamHI* en el extremo 5' de la secuencia de E7. El sitio *EcoRI* 3' se conservó. Los fragmentos resultantes se cortaron con *BamHI* y *EcoRI*, se purificaron y se ligaron en pCDNA3. Después de la secuenciación, los plásmidos se cortaron con *KpnI* y *BamHI* y se ligaron con las secuencias secretoras *KpnI/BamHI* relevantes. Las construcciones resultantes se secuenciaron y se nombraron IgkS1-50 a IgkS1-59 (véanse la TABLA 12 y las Figuras 12 y 14 para comparaciones de secuencias).

Construcciones de E7 secretadas con uso de codón predominantemente de tipo silvestre, modificadas para codones individuales

Se fabricaron construcciones que codificaban una forma secretada de E7 oncogénica (es decir, proteína E7 de tipo silvestre) por mutagénesis dirigida al sitio de los plásmidos que codificaban una forma secretada de E7 no oncogénica. Esto se hizo para construcciones para codones para los siguientes aminoácidos: Asp, Cys, Glu, Gln, Gly, Ile, Ser, Thr y Val.

La mutagénesis dirigida al sitio se realizó usando el kit de mutagénesis dirigida al sitio QuikChange II (Stratagene, La Jolla CA) y cebadores apropiados (Sigma) purificados por PAGE (electroforesis en gel de poliácridamida) de acuerdo con las instrucciones del fabricante. Las series de construcciones pcDNA-klgkX-nE7X se usaron como moldes para la mutagénesis (es decir, las construcciones IgkS1-13 a 24, IgkS1-27 a 29, IgkS1-34 a 43 y IgkS1-46 a 49). Los cebadores introdujeron las mutaciones G21D, G24C, G26E deseadas.

Las construcciones resultantes, IgkS2-13 a 24, IgkS2-27 a 29, IgkS2-34 a 43 y IgkS2-46 a 49 (véase la Tabla 8, SEQ ID NO: 1 a 29), tienen un uso de codón de tipo silvestre para la secuencia secretora Igk y la secuencia E7 con la excepción de que los codones para el aminoácido pertinente se cambiaron y codifican E7 oncogénica.

Construcciones enlazadoras

Se fabricaron construcciones que codifican la secuencia secretora Igk N terminal seguida de una secuencia enlazadora (XXGXGXX, donde X es el aminoácido pertinente para una construcción particular y G es glicina) y la proteína E7 para cada uno de los siguientes aminoácidos: Asn, Ala, Lys, Arg, Phe, His y Tyr.

Los fragmentos que consistían en la secuencia secretora Igk (con uso de codón de consenso de mamífero) y las secuencias enlazadoras se fabricaron por PCR usando Taq polimerasa y condiciones de ciclación convencionales, como se recomienda por el fabricante.

Los fragmentos se amplificaron a partir de pCDNA3-kcmIgk-cmE7 usando un cebador directo común (5'TTGAATAGGTACCGCCGCCACCATGGAGACCGACACCCTCC3'; SEQ ID NO: 90) que se hibridó con el sitio *KpnI*, la secuencia de Kozak y el comienzo de la secuencia secretora Igk. Los cebadores inversos fueron diferentes para cada construcción enlazadora y se hibridaron con el extremo de la secuencia secretora Igk (con uso de codón

de consenso de mamífero), introdujeron una nueva secuencia que codificó la secuencia enlazadora relevante y un sitio 3' *Bam*HI.

Los fragmentos se digirieron con *Kpn*II/*Bam*HI y se ligaron en pCDNA3-cmlgk-cmE7 cortado por *Kpn*II/*Bam*HI (es decir, la secuencia Kozak y la secuencia secretora se habían retirado del plásmido por digestión) para fabricar pCDNA3-cmlgk-enlazadorX-cmE7 (es decir, IgkS2-1 a 12, IgkS2-25 y 26, IgkS2-30 a 33 y IgkS2-44 y 45 como se ilustran en la Tabla 8, SEQ ID NO: 30 a 49).

Para los fragmentos de Asn también se ligaron en pCDNA3-Igk-nE7Asn1/2 cortado por *Kpn*II/*Bam*HI (es decir IgkS1-11 y 12) para fabricar pCDNA3-cmlgk-enlazadorN1/2-nE7Asn1/2 (es decir, IgkS2-11b y IgkS2-12b, véase la Tabla 12).

Expresión de proteína E7

Cultivo celular

Se cultivaron células CHO en DMEM (GIBCO de Invitrogen) que contenía suero bovino fetal (FBS) al 10 % (DKSH), penicilina, estreptomycin y glutamina (GIBCO de Invitrogen) a 37 °C y CO₂ al 5 %. Las células se sembraron en placas de 6 pocillos a 3 x 10⁵/pocillo, 24 horas antes de la transfección. Para cada transfección, se mezclaron 2 µg de ADN con 50 µl de OptiMEM (GIBCO de Invitrogen) y 4 µl de reactivo Plus (Invitrogen) y se incubaron a temperatura ambiente (TA) durante 30 min. Se añadió Lipofectamina (Invitrogen; 5 µl en 50 µl de OptiMEM) y los complejos se incubaron a TA durante 30 min. Las células se aclararon con OptiMEM, se añadieron 2 ml de OptiMEM a cada pocillo y después se añadieron los complejos. Las células se incubaron durante la noche a 37 °C y CO₂ al 5 %. A la mañana siguiente los complejos se retiraron y se añadieron 2 ml de DMEM recién preparado que contenía FBS al 2 % a cada pocillo.

Los sedimentos celulares y los sobrenadantes se recogieron aproximadamente 40 h después de la transfección. Los sedimentos celulares se resuspendieron en tampón de lisis (NP-40 al 0,1 %, aprotinina 2 µg/ml, leupeptina 1 µg/ml y PMSF 2 mM en PBS). Las transfecciones se realizaron por duplicado y se repitieron. También se realizaron transfecciones de control, con el vector vacío (pCDNA3).

Transferencia Western

Se realizaron transferencias Western de los sobrenadantes o lisados de células CHO de acuerdo con protocolos convencionales. Brevemente, esto implicó en primer lugar la separación de las muestras por electroforesis en gel de poliacrilamida (PAGE). Para los lisados de células, se cargaron 30 µg de proteína total para cada muestra. Para los sobrenadantes, se cargaron 30 µl de cada uno. Las muestras de proteína se hirvieron con tampón de carga de SDS-PAGE durante 10 minutos antes de cargarlas sobre geles de SDS-PAGE al 12 % y los geles se revelaron a 150-200 V durante aproximadamente 1 h.

Las proteínas separadas después se transfirieron desde los geles a una membrana de PVDF (100 V durante 1 h). Las membranas se bloquearon con leche desnatada al 5 % (en PBS/Tween 20 al 0,05 % (PBS-T)) durante 1 hora a temperatura ambiente y después se incubaron con el anticuerpo primario, Anticuerpo Monoclonal de Ratón VPH-16 E7 (Zymed Laboratories) a una concentración de 1:1000 en leche desnatada al 5 % (en PBS-T) durante la noche a 4 °C. Después del lavado de la membrana en PBS-T (10 minutos, 3 veces), se añadió anticuerpo secundario, anti-IgG de ratón (Sigma) en leche desnatada al 5 % y la membrana se incubó a temperatura ambiente durante 4 h. Las membranas se lavaron como anteriormente, se incubaron en una mezcla que contenía volúmenes iguales de solución A (4,425 ml de agua, 50 µl de luminol, 22 µl de ácido p-cumárico y 500 µl de Tris 1 M pH 8,5) y solución B (4,5 ml de agua, 3 µl H₂O₂ al 30 % y 500 µl de Tris 1 M, pH 8,5) durante 1 min y después se secaron y se envolvieron con una envoltura de plástico. La película se expuso a las transferencias durante diversos tiempos (1 min, 3 min o 10 min) y la película después se reveló.

Protocolos de inmunización por Pistola de genes

Purificación de plásmidos

Todos los plásmidos utilizados para la vacunación se cultivaron en la cepa de *Escherichia coli* DH5α y se purificaron usando el Kit Nucleobond Maxi (Machery-Nagel). La concentración de ADN se cuantificó espectrofotométricamente a 260 nm.

Preparación de cartuchos de ADN/oro

Se realizó el recubrimiento de partículas de oro con ADN plasmídico como se describe en el manual de instrucciones del Biorad Helios Gene Gun System usando una cantidad de carga de microvehículo (CCM) de 0,5 mg de oro/cartucho y una relación de carga de ADN de 2 µg de ADN/mg de oro. Esto dio lugar a 1 µg de ADN por cartucho preparado. Brevemente, se añadieron 50 µl de espermidina 0,05 M (Sigma) a 25 mg de partículas de oro 1,0 µm

(Bio-Rad) y la espermidina/oro se sometió a ultrasonidos durante 3 segundos. Después se añadieron 50 µg de ADN plasmídico, seguido de la adición gota a gota de 100 µl de CaCl₂ 1 M mientras se agitaba con formación de vórtice. La mezcla se dejó precipitar a temperatura ambiente durante 10 min, después se centrifugó para sedimentar el ADN/oro. El sedimento se lavó tres veces con etanol de calidad HPLC (Scharlau), antes de la resuspensión en etanol de calidad HPLC que contenía 0,5 mg/ml de polivinilpirrolidona (PVP) (Bio-Rad). Después, la suspensión de oro/plásmido se aplicó como recubrimiento a tubos Tefzel y cartuchos de 0,5 pulgadas (1,27 cm) preparados.

Inmunización de ratones por pistola de genes

Se inmunizaron grupos de 8 C57BL6/J hembras (de 6-8 semanas de edad) (ARC, WA or Monash Animal Services, VIC) el Día 0, el Día 21, el Día 42 y el Día 63 con el ADN pertinente. El día antes de cada inmunización se afeitó el abdomen de cada ratón y se aplicó crema depilatoria (Nair) durante 1 minuto. Se entregó ADN con la pistola de genes Helios (Biorad) usando una presión de 400 psi (2757,90 kPa). Los ratones recibieron 2 inyecciones en cada lado del abdomen, con 1 µg de ADN entregado por disparo. Se recogió suero mediante sangrado intraocular 2 días antes de la primera inmunización y 2 semanas después de cada inmunización posterior (Día 2, Día 35, Día 56 y Día 77).

ELISA para medir la respuesta inmunitaria a E7

Se usaron nueve péptidos que abarcaban la longitud completa de VPH16E7 (Frazer *et al.*, 1995) para medir la respuesta de anticuerpos a E7. Los péptidos se sintetizaron y se purificaron a una pureza del >70 % en Auspep (Melbourne). Se fabricaron los péptidos GF101 a 106 y GF108 a 109 descritos en Frazer *et al.* Nótese que en lugar de GF107, se usó GF107a: HYNIVTFCKCDSTLRL.

También se sintetizaron los péptidos GF102 D13G, GF103 D5G/C8G/E10G y GF104E2G, nombrados GF102n, GF103n y GF104n respectivamente. Estos péptidos se usaron para el ELISA cuando se midieron los anticuerpos frente a E7 no oncogénica, es decir, estos péptidos incorporan las mutaciones que se hicieron para hacer que la proteína E7 fuera no oncogénica.

Se recubrieron placas de microtitulación durante la noche con 50 µl de péptido E7 10 µg/ml por pocillo. Después del recubrimiento, las placas de microtitulación (Maxisorp, Nunc) se lavaron dos veces con PBS/Tween 20 al 0,05 % (PBS-T) y después se bloquearon durante dos horas a 37 °C con 100 µl de leche desnatada en polvo al 5 % en PBS-T. Después del bloqueo, las placas se lavaron tres veces con PBS-T y se añadieron 50 µl de suero de ratón a una dilución de 1 en 100 durante 2 horas a 37 °C. Todos los sueros se sometieron a ensayo en pocillos duplicados. Después, las placas se lavaron tres veces con PBS-T y se añadieron 50 µl conjugado de peroxidasa de rábano picante anti-IgG de ratón de oveja (Sigma) a una dilución de 1 en 1000. Después de 1 hora las placas se lavaron y se añadieron 50 µl de sustrato OPD. Se midió la absorbancia después de 30 min y la adición de 25 µl de HCl 2,5 M a 490 nm en un lector de placas Multiskan EX (Pathtech). Nótese los controles que se incluyeron: anticuerpo primario de control para un control positivo, anticuerpo secundario solamente y suero del día 0/suero de ratones sin inmunizar como controles negativos.

Las preferencias de respuesta inmunitaria de los codones determinadas a partir de estos experimentos se tabulan en la TABLA 1.

EJEMPLO 2

CONSTRUCCIÓN DE ADN DE HA DEL VIRUS DE LA GRIPE A (H5N1) MODIFICADO POR CODONES PARA CONFERIR UNA RESPUESTA INMUNITARIA POTENCIADA FRENTE A LA HA DE H5N1

La secuencia de nucleótidos de tipo silvestre del virus de la gripe A, gen HA para hemaglutinina (A/Hong Kong/213/03(H5N1), aislado por MDCK, aislado en huevo de pollo embrionado) se muestra en la SEQ ID NO: 50 y codifica la secuencia de aminoácidos que se muestra en la SEQ ID NO: 51. Se mutaron varios codones dentro de esa secuencia usando el método que se describe en el Ejemplo 1. Específicamente, el método implicó el reemplazo de codones de la secuencia de nucleótidos de tipo silvestre por codones sinónimos correspondientes que tenían preferencias de respuesta inmunitaria superiores con respecto a los codones que reemplazaron, como se representa en la Tabla 1. Una secuencia de nucleótidos modificada por codones ilustrativa que comprende codones de preferencia de respuesta inmunitaria alta se muestra en la SEQ ID NO: 52.

EJEMPLO 3

CONSTRUCCIÓN DE ADN DEL VIRUS DE LA GRIPE A (H3N1) MODIFICADO POR CODONES PARA CONFERIR UNA RESPUESTA INMUNITARIA POTENCIADA FRENTE A LA HA DE H3N1

La secuencia de nucleótidos de tipo silvestre del virus de la gripe A, gen HA para hemaglutinina (A/cerdo/Corea/PZ72-1/2006(H3N1)) se muestra en la SEQ ID NO: 53 y codifica la secuencia de aminoácidos que se muestra en la SEQ ID NO: 54. Específicamente, el método implicó el reemplazo de codones de la secuencia de

nucleótidos de tipo silvestre por codones sinónimos correspondientes que tenían preferencias de respuesta inmunitaria superiores con respecto a los codones que reemplazaron, como se representa en la Tabla 1. Una secuencia de nucleótidos modificada por codones ilustrativa que comprende codones de preferencia de respuesta inmunitaria alta se muestra en la SEQ ID NO: 55.

EJEMPLO 4

CONSTRUCCIÓN DE ADN DE NA DEL VIRUS DE LA GRIPE A (H5N1) MODIFICADO POR CODONES PARA CONFERIR UNA RESPUESTA INMUNITARIA POTENCIADA FRENTE A LA NA DE H5N1

La secuencia de nucleótidos de tipo silvestre del virus de la gripe A, gen NA para neuraminidasa (A/Hong Kong/213/03(H5N1), gen de neuraminidasa NA, aislado por MDCK, aislado en huevo de pollo embrionado) se muestra en la SEQ ID NO: 50 y codifica la secuencia de aminoácidos que se muestra en la SEQ ID NO: 57. Se mutaron varios codones dentro de esa secuencia usando el método que se describe en el Ejemplo 1. Específicamente, el método implicó el reemplazo de codones de la secuencia de nucleótidos de tipo silvestre por codones sinónimos correspondientes que tenían preferencias de respuesta inmunitaria superiores con respecto a los codones que reemplazaron, como se representa en la Tabla 1. Una secuencia de nucleótidos modificada por codones ilustrativa que comprende codones de preferencia de respuesta inmunitaria alta se muestra en la SEQ ID NO: 58.

EJEMPLO 5

CONSTRUCCIÓN DE ADN DE NA DEL VIRUS DE LA GRIPE A (H3N1) MODIFICADO POR CODONES PARA CONFERIR UNA RESPUESTA INMUNITARIA POTENCIADA FRENTE A LA NA DE H3N1

La secuencia de nucleótidos de tipo silvestre del virus de la gripe A, gen NA para neuraminidasa (A/cerdo/MI/PU243/04(H3N1)) se muestra en la SEQ ID NO: 59 y codifica la secuencia de aminoácidos que se muestra en la SEQ ID NO: 60. Se mutaron varios codones dentro de esa secuencia usando el método que se describe en el Ejemplo 1. Específicamente, el método implicó el reemplazo de codones de la secuencia de nucleótidos de tipo silvestre por codones sinónimos correspondientes que tenían preferencias de respuesta inmunitaria superiores con respecto a los codones que reemplazaron, como se representa en la Tabla 1. Una secuencia de nucleótidos modificada por codones ilustrativa que comprende codones de preferencia de respuesta inmunitaria alta se muestra en la SEQ ID NO: 61.

EJEMPLO 6

CONSTRUCCIÓN DE ADN DE E1 DEL VIRUS DE LA HEPATITIS C (1AH77) MODIFICADO POR CODONES PARA CONFERIR UNA RESPUESTA INMUNITARIA POTENCIADA FRENTE AL E1 DEL VHC (1AH77)

La secuencia de nucleótidos de tipo silvestre de E1 del virus de la hepatitis C, (serotipo 1A, aislado H77, de la secuencia de nucleótidos de poliproteína AF009606) se muestra en la SEQ ID NO: 62 y codifica la secuencia de aminoácidos (NP 751920) que se muestra en la SEQ ID NO: 63. Se mutaron varios codones dentro de esa secuencia usando el método que se describe en el Ejemplo 1. Específicamente, el método implicó el reemplazo de codones de la secuencia de nucleótidos de tipo silvestre por codones sinónimos correspondientes que tenían preferencias de respuesta inmunitaria superiores con respecto a los codones que reemplazaron, como se representa en la Tabla 1. Una secuencia de nucleótidos modificada por codones ilustrativa que comprende codones de preferencia de respuesta inmunitaria alta se muestra en la SEQ ID NO: 64.

EJEMPLO 7

CONSTRUCCIÓN DE ADN DE E2 DEL VIRUS DE LA HEPATITIS C (1AH77) MODIFICADO POR CODONES PARA CONFERIR UNA RESPUESTA INMUNITARIA POTENCIADA FRENTE AL E2 DEL VHC (1AH77)

La secuencia de nucleótidos de tipo silvestre de E2 del virus de la hepatitis C, (serotipo 1A, aislado H77, de la secuencia de nucleótidos de poliproteína AF009606) se muestra en la SEQ ID NO: 65 y codifica la secuencia de aminoácidos (NP 751921) que se muestra en la SEQ ID NO: 66. Se mutaron varios codones dentro de esa secuencia usando el método que se describe en el Ejemplo 1. Específicamente, el método implicó el reemplazo de codones de la secuencia de nucleótidos de tipo silvestre por codones sinónimos correspondientes que tenían preferencias de respuesta inmunitaria superiores con respecto a los codones que reemplazaron, como se representa en la Tabla 1. Una secuencia de nucleótidos modificada por codones ilustrativa que comprende codones de preferencia de respuesta inmunitaria alta se muestra en la SEQ ID NO: 67.

EJEMPLO 8

CONSTRUCCIÓN DE ADN DE VIRUS DE EPSTEIN-BARR DE TIPO 1 GP350 MODIFICADO POR CODONES PARA CONFERIR UNA RESPUESTA INMUNITARIA POTENCIADA FRENTE AL VEB DE TIPO 1 GP350

La secuencia de nucleótidos de tipo silvestre del virus de Epstein-Barr, VEB de tipo 1 gp350 (gen BLLF1, cadena 77142-79865) se muestra en la SEQ ID NO: 68 y codifica la secuencia de aminoácidos (CAD53417) que se muestra en la SEQ ID NO: 69. Se mutaron varios codones dentro de esa secuencia usando el método que se describe en el Ejemplo 1. Específicamente, el método implicó el reemplazo de codones de la secuencia de nucleótidos de tipo silvestre por codones sinónimos correspondientes que tenían preferencias de respuesta inmunitaria superiores con respecto a los codones que reemplazaron, como se representa en la Tabla 1. Una secuencia de nucleótidos modificada por codones ilustrativa que comprende codones de preferencia de respuesta inmunitaria alta se muestra en la SEQ ID NO: 70.

EJEMPLO 9

CONSTRUCCIÓN DE ADN DE VIRUS DE EPSTEIN-BARR DE TIPO 2 GP350 MODIFICADO POR CODONES PARA CONFERIR UNA RESPUESTA INMUNITARIA POTENCIADA FRENTE AL VEB DE TIPO 2 GP350

La secuencia de nucleótidos de tipo silvestre del virus de Epstein-Barr, VEB de tipo 2 gp350 (gen BLLF1, cadena 77267-29936) se muestra en la SEQ ID NO: 71 y codifica la secuencia de aminoácidos (YP 0011294462) que se muestra en la SEQ ID NO: 72. Se mutaron varios codones dentro de esa secuencia usando el método que se describe en el Ejemplo 1. Específicamente, el método implicó el reemplazo de codones de la secuencia de nucleótidos de tipo silvestre por codones sinónimos correspondientes que tenían preferencias de respuesta inmunitaria superiores con respecto a los codones que reemplazaron, como se representa en la Tabla 1. Una secuencia de nucleótidos modificada por codones ilustrativa que comprende codones de preferencia de respuesta inmunitaria alta se muestra en la SEQ ID NO: 73.

EJEMPLO 10

CONSTRUCCIÓN DE ADN DE GLUCOPROTEÍNA B DEL VIRUS DEL HERPES SIMPLE 2 MODIFICADO POR CODONES PARA CONFERIR UNA RESPUESTA INMUNITARIA POTENCIADA FRENTE A LA GLUCOPROTEÍNA B DEL VHS-2

La secuencia de nucleótidos de tipo silvestre del Virus del Herpes Simple 2, glucoproteína B cepa HG52 (genoma cepa NC 001798) se muestra en la SEQ ID NO: 74 y codifica la secuencia de aminoácidos (CAB06752) que se muestra en la SEQ ID NO: 75. Se mutaron varios codones dentro de esa secuencia usando el método que se describe en el Ejemplo 1. Específicamente, el método implicó el reemplazo de codones de la secuencia de nucleótidos de tipo silvestre por codones sinónimos correspondientes que tenían preferencias de respuesta inmunitaria superiores con respecto a los codones que reemplazaron, como se representa en la Tabla 1. Una secuencia de nucleótidos modificada por codones ilustrativa que comprende codones de preferencia de respuesta inmunitaria alta se muestra en la SEQ ID NO: 76.

EJEMPLO 11

CONSTRUCCIÓN DE ADN DE GLUCOPROTEÍNA D DEL VIRUS DEL HERPES SIMPLE 2 MODIFICADO POR CODONES PARA CONFERIR UNA RESPUESTA INMUNITARIA POTENCIADA FRENTE A LA GLUCOPROTEÍNA D DEL VHS-2

La secuencia de nucleótidos de tipo silvestre del Virus del Herpes Simple 2, glucoproteína D cepa HG52 (genoma cepa NC 001798) se muestra en la SEQ ID NO: 77 y codifica la secuencia de aminoácidos (NP 044536) que se muestra en la SEQ ID NO: 78. Se mutaron varios codones dentro de esa secuencia usando el método que se describe en el Ejemplo 1. Específicamente, el método implicó el reemplazo de codones de la secuencia de nucleótidos de tipo silvestre por codones sinónimos correspondientes que tenían preferencias de respuesta inmunitaria superiores con respecto a los codones que reemplazaron, como se representa en la Tabla 1. Una secuencia de nucleótidos modificada por codones ilustrativa que comprende codones de preferencia de respuesta inmunitaria alta se muestra en la SEQ ID NO: 79.

EJEMPLO 12

CONSTRUCCIONES OPTIMIZADAS DE E7 y VHS-2

Diseño y síntesis de construcciones de E7 óptimas y menos óptimas

Una construcción de E7 desoptimizada (W) y tres optimizadas (O1-O3) se diseñaron y se fabricaron usando las preferencias de codón que se resumen en la Tabla 1 ("la tabla Inmune Coricode inmunitaria"). Los codones menos favorables se usaron para la construcción W. Para la primera construcción optimizada, O1, cuya secuencia se muestra en la SEQ ID NO: 81, todos los codones se modificaron a aquellos codones determinados como más óptimos. O2, cuya secuencia se muestra en la SEQ ID NO: 82, es una construcción optimizada alternativa que implicó cambiando toda la Ala a GCT; Arg CGG y AGG a CGA y AGA, respectivamente; Glu a GAA; Gly a GGA; Ile al ATC; toda la Leu a CTG; Phe a TTT, Pro a CCT o CCC, Ser a TCG, Thr a ACG; y toda la Val excepto GTG a

GTC. Las modificaciones 02 evitaron, con la excepción de Leu y Ile, cambiar codones a codones preferidos de consenso de mamíferos. Para 03, cuya secuencia se muestra en la SEQ ID NO: 83, solamente se modificaron determinados aminoácidos para los que se observaron diferencias particularmente distintas entre codones y para los que el codón o codones óptimos tampoco fuera un codón preferido de consenso de mamífero. En particular, en 03 todos los codones de Gly, Leu, Pro, Ser y Thr no preferida se cambiaron a GGA, CTC, CCT, TCG y ACG, respectivamente, y en los que ya se había usado un codón preferido éste no se alteró. Los codones para otros aminoácidos en 03 no se modificaron.

Respuestas humoral y celular a la inmunización biolística con las construcciones de E7 óptimas y menos óptimas

Como puede observarse en la Figura 18 (a) las tres construcciones optimizadas (O1 a 03) dieron lugar a respuestas de anticuerpos significativamente mayores que la construcción de tipo silvestre medida tanto por el ELISA de péptido como el ELISA de una proteína GST-E7. Las amplitudes de la respuesta no fueron estadísticamente diferentes entre las tres construcciones optimizadas. La construcción desoptimizada, W, cuya secuencia se muestra en la SEQ ID NO: 84, dio una respuesta de anticuerpos muy baja, pareciendo ligeramente inferior pero no estadísticamente diferente de la construcción de uso de codón (UC) de tipo silvestre (ts), cuya secuencia se muestra en la SEQ ID NO: 80. A partir de los experimentos de ELISPOT de IFN- γ , un ejemplo representativo de los cuales se muestra en la Figura 18, parece que las preferencias de codón para maximizar la respuesta de anticuerpos son similares a las necesarias para maximizar la respuesta de linfocitos T: la construcción desoptimizada W no consiguió proporcionar una respuesta medible en el ensayo ELISPOT de IFN- γ y dos de las construcciones optimizadas (02 y 03) proporcionaron respuestas estadísticamente significativamente mayores que la construcción de UC de tipo silvestre. Durante las tres repeticiones, las respuestas a 02 y 03 no fueron estadísticamente diferentes entre sí. Inesperadamente y en contraste con la tendencia de anticuerpos, en dos de los tres experimentos repetidos O1 proporcionó una respuesta celular similar a la construcción de UC ts, que era menor que la conseguida por las construcciones 02 o 03.

Respuestas humoral y celular a la inmunización por inyección intradérmica con las construcciones de E7 óptimas y menos óptimas

Las respuestas humorales y celulares de ratones a las construcciones optimizadas, de UC de tipo silvestre y desoptimizadas entregadas por inyección intradérmica también se midieron y los resultados se resumen en la Figura 19. En general, se observaron tendencias similares tanto para la inyección intradérmica como para la entrega biolística.

A partir del ELISA de la proteína E7, es evidente que las tres construcciones optimizadas, O1-O3, fueron todas significativamente mejores en la generación de anticuerpos que la construcción de tipo silvestre y que la construcción desoptimizada proporcionó una respuesta de anticuerpos muy baja similar a la de tipo silvestre. Las construcciones optimizadas todas dieron lugar a significativamente más manchas en el ELISPOT de IFN- γ que la construcción de tipo silvestre y la construcción desoptimizada no consiguió dar lugar a una respuesta medible.

Las amplitudes de las respuestas de anticuerpos a la inmunización por pistola de genes fueron más grandes que para las vacunas entregadas por vía intradérmica (ID), a pesar de que la inmunización por vía ID entrega más de cinco veces la dosis.

Diseño y síntesis de construcciones de VHS-2 óptimas y menos óptimas

Se prepararon tres construcciones optimizadas (O1-O3; cuyas secuencias se muestran en las SEQ ID NO: 86-88, respectivamente) y una construcción desoptimizada (W; cuya secuencia se muestra en la SEQ ID NO: 88) que codifican la glucoproteína D de longitud completa del virus del Herpes Simple 2 (gD2). También se fabricó una construcción de control de pCDNA3-gD2 con UC ts. El UC de tipo silvestre, cuya secuencia se muestra en la SEQ ID NO: 85, está cerca del UC de CM.

Respuestas humorales a la inmunización biolística e intradérmica con las construcciones de gD2 óptimas y menos óptimas

Se inmunizaron ratones C57B1/6 en dos grupos (8 ratones/construcción; se usaron inyección intradérmica (ID) y entrega por pistola de genes) usando el mismo protocolo de inmunización que para las construcciones de E7.

El grupo 1 incluyó pCDNA3-gD2 y pCDNA3-gD2 O1. El grupo 2 incluyó pCDNA3-gD2, pCDNA3-gD2 02, pCDNA3-gD2 03 y pCDNA3-gD2 W.

Las respuestas de anticuerpos se midieron mediante un ELISA usando placas recubiertas con sobrenadante de células CHO que contenía gD2 marcado con His en el extremo C terminal y truncado. El truncamiento es en el resto de aminoácido 331 y elimina la región transmembrana que da como resultado que la proteína se secrete en el medio. Se usaron como control placas de ELISA de control recubiertas con sobrenadante de células CHO transfectadas con vector vacío.

Para ambas vías de entrega biolística y por inyección intradérmica se descubrió que las tres construcciones optimizadas generaron niveles similares de anticuerpos que la construcción de gD2 de UC de ts (Figura 20). La construcción desoptimizada, W gD2, fue muy mala en la generación de anticuerpos, en particular cuando se entregó mediante inyección intradérmica. Los dos métodos de entrega dieron como resultado niveles de anticuerpos similares.

Hasta la fecha, no existen vacunas de ADN en el mercado para el tratamiento o la prevención de enfermedades en los seres humanos. Existe una necesidad de maximizar las respuestas inmunitarias generadas por las vacunas de ADN y la presente invención desvela maneras de potenciar la eficacia de las vacunas de ADN mediante el uso de codones que tienen una preferencia mayor para la producción de una respuesta inmunitaria.

El estudio descrito en este ejemplo ha validado la tabla Inmune Coricode mediante su aplicación a la optimización o desoptimización de los genes de E7 de VPH16 y de la glucoproteína D (gD2) de VHS-2 y la demostración de que esto potencia o reduce, respectivamente, la respuesta de anticuerpos o celular a la entrega biolística de estos genes a mamíferos tales como ratones.

Material y Métodos

Ensayo ELISPOT

Para los ELISPOT de IFN- γ , los ratones se inmunizaron dos veces, en los días 0 y 21 y los bazos se recogieron 3 semanas después de la segunda inmunización.

Protocolo de inyección intradérmica

El momento y la frecuencia de las inmunizaciones por inyección intradérmica fueron los mismos que para la inmunización por pistola de genes. En cada inmunización se inyectaron 5 μ g de ADN por oído, es decir, se administraron un total de 10 μ g por inmunización por ratón. La eliminación del pelo antes de la inmunización no fue necesaria. El momento de las hemorragias y el momento de la recolección del bazo fueron los mismos que para los ratones inmunizados por pistola de genes.

ELISA de GST-E7

El ELISA de GST-E7 se realizó de la misma manera que el ELISA de péptidos con la excepción de que las placas se recubrieron durante la noche con 50 μ l de proteína E7 10 μ g/ml marcada con GST (amablemente proporcionada por el grupo de Frazer del Instituto Diamantina, Universidad de Queensland, Brisbane).

ELISA de gD de VHS-2

Este ELISA se realizó de la misma manera que los ELISA de E7 con la excepción de que las placas se recubrieron con sobrenadante de células CHO transfectadas con un vector que codifica proteína gD2 marcada con His en el extremo C terminal y truncada. También se usaron placas de control recubiertas con sobrenadante de células CHO transfectadas con vector vacío.

Detección de respuestas específicas del VPH

Para la detección de respuestas específicas del VPH, se recubrieron placas ELISPOT de filtro de 96 pocillos (Millipore) durante la noche con proteína E7 marcada con GST de VPH 10 μ g/ml en NaHCO_3 0,1 M. Para la detección de células secretoras de IgG totales, se recubrieron placas ELISPOT de filtro de 96 pocillos durante la noche con anti-Ig de ratón de cabra 2 μ g/ml (Sigma) en PBS sin MgCl_2 ni CaCl_2 . Después del recubrimiento, las placas se lavaron una vez con DMEM completo sin FCS y después se bloquearon con DMEM completo complementado con FCS al 10 % durante una hora a 37 °C. Las células de bazo de ratón cultivadas se lavaron y se añadieron a placas ELISPOT a 10^6 células/100 μ l. Para la detección de células B de memoria específicas del VPH, las placas se incubaron durante la noche a 37 °C y para la medición de células IgG totales, las placas se incubaron durante 1 hora a 37 °C. Para la detección, se usó anti-IgG de ratón biotinilado de cabra (Sigma) en PBS-T/FCS al 1 %, seguido de avidina conjugada con HRP 5 μ g/ml (Pierce) y se reveló usando 3-amino-9-etilcarbazol (Sigma). Las placas reveladas se contaron usando un contador de placas ELISPOT automatizado.

ELISPOT de IFN- γ de E7

Se recubrieron placas de filtro de 96 pocillos (Millipore) durante la noche con 4 μ g/ml de anticuerpo monoclonal (AN18; Mabtech). Después del recubrimiento, las placas se lavaron una vez con RPMI completo y se bloquearon durante 2 horas con RPMI completo con suero de ternera fetal al 10 % (FCS; CSL Ltd). Los bazos de ratón se convirtieron en suspensiones de células individuales y se trataron con tampón de lisis de ACK, se lavaron y se resuspendieron a una concentración de 10^7 células/ml. Se añadieron células de bazo (10^6 /pocillo) a cada pocillo, seguido de la adición de RPMI completo complementado con hIL-2 recombinante (ProSpec-Tany TechnoGene Ltd) y

péptido hasta una concentración final de 10 UI/pocillo y 1 µg/ml, respectivamente. Se añadió medio que contenía hIL-2 sin péptido a los pocillos de control. Las placas se incubaron durante aproximadamente 18 horas a 37 °C en CO₂ al 5-8 %.

Después de la incubación durante la noche, las células se lisaron mediante aclarado de las placas en agua corriente y después se lavaron seis veces en PBS/Tween 20 al 0,05 % (PBS-T). Para la detección, se añadió mAb de detección biotinilado (R4-6A2; Mabtech) en PBS-T/FCS al 2 %, seguido de peroxidasa de rábano picante (HRP) conjugada con estreptavidina y DAB (Sigma). Las placas reveladas se contaron usando un contador de placas ELISPOT automatizado.

BIBLIOGRAFÍA

Ausubel, F.M. (Ed.) 2007. *Current Protocols in Molecular Biology*. Ebook (<http://www.mrw.interscience.wiley.com/emrw/9780471142720%p/cpmb/toc>).

Edmonds, C. y Vousden, K.H. (1989). *A point mutational analysis of human papillomavirus type 16 E7 protein*. *Journal of Virology*. 63: 2650-2656.

Frazer, I.H., Leippe, D.M., Dunn, L.A., Leim, A., Tindle, R.W., Fernando, G.J., Phelps, W.C. y Lambert, P.F. (1995). *Immunological responses in human papillomavirus 16 E6/E7 transgenic mice to E7 protein correlate with the presence of skin disease*. *Cancer Research*. 55: 2635-2639.

Heck, D.V., Yee, C.L., Howley, P.M. y Munger, K. (1992). *Efficiency of binding the retinoblastoma protein correlates with the transforming capacity of the E7 oncoproteins of the human papillomaviruses*. *PNAS* 89: 4442-4446.

Liu, W.J., Gao, F., Zhao, K.N., Zhao, W., Fernando, G.J., Thomas, R. y Frazer, I.H. (2002). *Codon modified human papillomavirus type 16 E7 DNA vaccine enhances cytotoxic T-lymphocyte induction and anti-tumour activity*. *Virology* 301: 43-52.

Smith, H.O., Hutchison III, C.A., Pfannkoch, C. y Venter, J.C. (2003). *Generating a synthetic genome by whole genome assembly: IX174 bacteriophage from synthetic oligonucleotides*. *PNAS*. 100 (26): 15440-15445.

LISTADO DE SECUENCIAS

<110> The University of Queensland (excepto EE.UU.) Frazer, Ian H (solo EE.UU.)

<120> SISTEMA DE EXPRESIÓN PARA LA MODULACIÓN DE UNA RESPUESTA INMUNITARIA

<130> 40130500

<140> Sin asignar
<141> 02-10-2008

<150> US 60/980.145
<151> 15-10-2007

<160> 91

<170> PatentIn versión 3.5

<210> 1
<211> 387
<212> ADN

<213> Secuencia artificial

<220>
<223> Secuencia plasmídica

<400> 1

ES 2 662 030 T3

```

ggtagcgccg ccaccatgga gacagataca ctctgctat gggtagctgt gctctgggtt    60
ccagggtcca ctgggtgatgg atccatgcat ggagatacac ctacattgca tgaatatatg    120
ttagatttgc aaccagagac aactgatctc tactgttatg agcaattaaa tgatagctca    180
gaggaggagg atgaaataga tgggtccagct ggacaagcag aaccggatag agcccattac    240
aatattgtaa ccttttggtg caagtgtgat tctacgcttc ggttgtgctg acaaagcaca    300
cacgtagata ttctactttt ggaagatctg ttaatgggca cactaggaat tgtgtgcccc    360
atctgctctc agaagcccta agaattc                                     387

```

<210> 2
 <211> 387
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Secuencia plasmídica

<400> 2

```

ggtagcgccg ccaccatgga gacagacaca ctctgctat gggtagctgt gctctgggtt    60
ccagggtcca ctgggtgacgg atccatgcat ggagacacac ctacattgca tgaatatatg    120
ttagacttgc aaccagagac aactgacctc tactgttatg agcaattaaa tgacagctca    180
gaggaggagg acgaaataga cgggtccagct ggacaagcag aaccggacag agcccattac    240
aatattgtaa ccttttggtg caagtgtgac tctacgcttc ggttgtgctg acaaagcaca    300
cacgtagaca ttctactttt ggaagacctg ttaatgggca cactaggaat tgtgtgcccc    360
atctgctctc agaagcccta agaattc                                     387

```

<210> 3
 <211> 387
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Secuencia plasmídica

<400> 3

```

ggtagcgccg ccaccatgga gacagacaca ctctgctat gggtagctgt gctctgggtt    60
ccagggtcca ctgggtgacgg atccatgcat ggagatacac ctacattgca tgaatatatg    120
ttagatttgc aaccagagac aactgatctc tactgttatg agcaattaaa tgacagctca    180
gaggaggagg atgaaataga tgggtccagct ggacaagcag aaccggacag agcccattac    240
aatattgtaa ccttttggtg taagtgtgac tctacgcttc ggttgtgtgt acaaagcaca    300
cacgtagaca ttctactttt ggaagacctg ttaatgggca cactaggaat tgtgtgtccc    360
atctgttctc agaagcccta agaattc                                     387

```

<210> 4
 <211> 387
 <212> ADN
 <213> Secuencia artificial

<220>
<223> Secuencia plasmídica

<400> 4

5

```

ggtagcgccg ccaccatgga gacagacaca ctctgctat gggtagctgt gctctgggtt      60
ccagggttcca ctgggtgacgg atccatgcat ggagatacac ctacattgca tgaatatatg      120
ttagatttgc aaccagagac aactgatctc tactgctatg agcaattaaa tgacagctca      180
gaggaggagg atgaaataga tgggtccagct ggacaagcag aaccggacag agcccattac      240
aatattgtaa ccttttgctg caagtgcgac tctacgcttc ggttggtcgt acaaagcaca      300
cacgtagaca ttcgtacttt ggaagacctg ttaatgggca cactaggaat tgtgtgcccc      360
atctgctctc agaagcccta agaattc      387

```

<210> 5
<211> 387
<212> ADN
<213> Secuencia artificial

10

<220>
<223> Secuencia plasmídica

15

<400> 5

```

ggtagcgccg ccaccatgga gacagacaca ctctgctat gggtagctgt gctctgggtt      60
ccagggttcca ctgggtgacgg atccatgcat ggagatacac ctacattgca tgagtatatg      120
ttagatttgc aaccagagac aactgatctc tactgttatg agcaattaaa tgacagctca      180
gaggaggagg atgagataga tgggtccagct ggacaagcag agccggacag agcccattac      240
aatattgtaa ccttttgctg caagtgtgac tctacgcttc ggttggtcgt acaaagcaca      300
cacgtagaca ttcgtacttt ggaggacctg ttaatgggca cactaggaat tgtgtgcccc      360
atctgctctc agaagcccta agaattc      387

```

20

<210> 6
<211> 387
<212> ADN
<213> Secuencia artificial

25

<220>
<223> Secuencia plasmídica

<400> 6

```

ggtagcgccg ccaccatgga aacagacaca ctctgctat gggtagctgt gctctgggtt      60
ccagggttcca ctgggtgacgg atccatgcat ggagatacac ctacattgca tgaatatatg      120
ttagatttgc aaccagaaac aactgatctc tactgttatg aacaattaaa tgacagctca      180
gaagaagaag atgaaataga tgggtccagct ggacaagcag aaccggacag agcccattac      240
aatattgtaa ccttttgctg caagtgtgac tctacgcttc ggttggtcgt acaaagcaca      300
cacgtagaca ttcgtacttt ggaagacctg ttaatgggca cactaggaat tgtgtgcccc      360
atctgctctc agaagcccta agaattc      387

```

30

ES 2 662 030 T3

5	<210> 7 <211> 387 <212> ADN <213> Secuencia artificial	
	<220> <223> Secuencia plasmídica	
10	<400> 7	
	ggtaccgccg ccaccatgga gacagacaca ctcctgctat ggggtactgct gctctggggtt	60
	ccagggttcca ctgggtgacgg atccatgcat ggagatacac ctacattgca tgaatatatg	120
	ttagatttgc agccagagac aactgatctc tactgttatg agcagttaa tgacagctca	180
	gaggaggagg atgaaataga tgggtccagct ggacaggcag aaccggacag agcccattac	240
	aatattgtaa ccttttggtg caagtgtgac tctacgcttc ggttgtgctg acagagcaca	300
	cacgtagaca ttcgtacttt ggaagacctg ttaatgggca cactaggaat tgtgtgcccc	360
	atctgctctc agaagcccta agaattc	387
15	<210> 8 <211> 387 <212> ADN <213> Secuencia artificial	
20	<220> <223> Secuencia plasmídica	
	<400> 8	
	ggtaccgccg ccaccatgga gacagacaca ctcctgctat ggggtactgct gctctggggtt	60
	ccagggttcca ctgggtgacgg atccatgcat ggagatacac ctacattgca tgaatatatg	120
	ttagatttgc aaccagagac aactgatctc tactgttatg agcaattaaa tgacagctca	180
	gaggaggagg atgaaataga tgggtccagct ggacaagcag aaccggacag agcccattac	240
	aatattgtaa ccttttggtg caagtgtgac tctacgcttc ggttgtgctg acaaagcaca	300
	cacgtagaca ttcgtacttt ggaagacctg ttaatgggca cactaggaat tgtgtgcccc	360
25	atctgctctc aaaagcccta agaattc	387
30	<210> 9 <211> 387 <212> ADN <213> Secuencia artificial	
	<220> <223> Secuencia plasmídica	
35	<400> 9	

ES 2 662 030 T3

```

ggtagcgccg ccaccatgga gacagacaca ctctgtctat gggtagctgt gctctgggtt      60
ccaggggtcca ctggggacgg atccatgcat ggggatacac ctacattgca tgaatatatg      120
ttagatttgc aaccagagac aactgatctc tactgttatg agcaattaaa tgacagctca      180
gaggaggagg atgaaataga tgggccagct gggcaagcag aaccggacag agcccattac      240
aatattgtaa ccttttggtg caagtgtgac tctacgcttc ggttgtgcgt acaaagcaca      300
cacgtagaca ttcgtacttt ggaagacctg ttaatgggga cactagggat tgtgtgcccc      360
atctgctctc agaagcccta agaattc                                          387

```

<210> 10
 <211> 387
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Secuencia plasmídica

<400> 10

```

ggtagcgccg ccaccatgga gacagacaca ctctgtctat gggtagctgt gctctgggtt      60
ccaggttcca ctgggacgg atccatgcat ggagatacac ctacattgca tgaatatatg      120
ttagatttgc aaccagagac aactgatctc tactgttatg agcaattaaa tgacagctca      180
gaggaggagg atgaaataga tggaccagct ggacaagcag aaccggacag agcccattac      240
aatattgtaa ccttttggtg caagtgtgac tctacgcttc ggttgtgcgt acaaagcaca      300
cacgtagaca ttcgtacttt ggaagacctg ttaatgggaa cactaggaat tgtgtgcccc      360
atctgctctc agaagcccta agaattc                                          387

```

<210> 11
 <211> 387
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Secuencia plasmídica

<400> 11

```

ggtagcgccg ccaccatgga gacagacaca ctctgtctat gggtagctgt gctctgggtt      60
ccaggttcca ctggtgacgg atccatgcat ggtgatacac ctacattgca tgaatatatg      120
ttagatttgc aaccagagac aactgatctc tactgttatg agcaattaaa tgacagctca      180
gaggaggagg atgaaataga tggccagct ggtcaagcag aaccggacag agcccattac      240
aatattgtaa ccttttggtg caagtgtgac tctacgcttc ggttgtgcgt acaaagcaca      300
cacgtagaca ttcgtacttt ggaagacctg ttaatgggta cactaggtat tgtgtgcccc      360
atctgctctc agaagcccta agaattc                                          387

```

<210> 12
 <211> 387
 <212> ADN
 <213> Secuencia artificial

<220>
<223> Secuencia plasmídica

<400> 12

5

```

ggtagcgccg ccaccatgga gacagacaca ctctgctat gggtagctgt gctctgggtt      60
ccagggtcca ctggcgacgg atccatgcat ggagatacac ctacattgca tgaatatatg      120
ttagatttgc aaccagagac aactgatctc tactgttatg agcaattaaa tgacagctca      180
gaggaggagg atgaaataga tggccagct ggacaagcag aaccggacag agcccattac      240
aatattgtaa ccttttggtg caagtgtgac tctacgcttc ggttgtgcgt acaaagcaca      300
cacgtagaca ttctactttt ggaagacctg ttaatgggca cactaggcat tgtgtgcccc      360
atctgctctc agaagcccta agaattc                                          387

```

<210> 13
<211> 387
<212> ADN
<213> Secuencia artificial

10

<220>
<223> Secuencia plasmídica

15

<400> 13

```

ggtagcgccg ccaccatgga gacagacaca ctctgctat gggtagctgt gctctgggtt      60
ccagggtcca ctggcgacgg atccatgcat ggagatacac ctacattgca tgaatatatg      120
ttagatttgc aaccagagac aactgatctc tactgttatg agcaattaaa tgacagctca      180
gaggaggagg atgaaataga tggccagct ggacaagcag aaccggacag agcccattac      240
aatatagtaa ccttttggtg caagtgtgac tctacgcttc ggttgtgcgt acaaagcaca      300
cacgtagaca tacgtacttt ggaagacctg ttaatgggca cactaggaat agtgtgcccc      360
atatgctctc agaagcccta agaattc                                          387

```

20

<210> 14
<211> 387
<212> ADN
<213> Secuencia artificial

25

<220>
<223> Secuencia plasmídica

<400> 14

```

ggtagcgccg ccaccatgga gacagacaca ctctgctat gggtagctgt gctctgggtt      60
ccagggtcca ctggcgacgg atccatgcat ggagatacac ctacattgca tgaatatatg      120
ttagatttgc aaccagagac aactgatctc tactgttatg agcaattaaa tgacagctca      180
gaggaggagg atgaaattga tggccagct ggacaagcag aaccggacag agcccattac      240
aatattgtaa ccttttggtg caagtgtgac tctacgcttc ggttgtgcgt acaaagcaca      300
cacgtagaca ttctactttt ggaagacctg ttaatgggca cactaggaat tgtgtgcccc      360
atttgctctc agaagcccta agaattc                                          387

```

30

ES 2 662 030 T3

5	<210> 15		
	<211> 387		
	<212> ADN		
	<213> Secuencia artificial		
	<220>		
10	<223> Secuencia plasmídica		
	<400> 15		
	ggtaccgccg ccaccatgga gacagacaca ctctgctat gggactgct gctctgggtt	60	
	ccagggtcca ctggtgacgg atccatgcat ggagatacac ctacattgca tgaatatatg	120	
	ttagatttgc aaccagagac aactgatctc tactgttatg agcaattaaa tgacagctca	180	
15	gaggaggagg atgaaatcga tggccagct ggacaagcag aaccggacag agcccattac	240	
	aatatcgtaa ccttttgttg caagtgtgac tctacgcttc ggttgtgctg acaaagcaca	300	
	cacgtagaca tccgtacttt ggaagacctg ttaatgggca cactaggaat cgtgtgcccc	360	
	atctgctctc agaagcccta agaattc	387	
	<210> 16		
20	<211> 387		
	<212> ADN		
	<213> Secuencia artificial		
	<220>		
	<223> Secuencia plasmídica		
25	<400> 16		
	ggtaccgccg ccaccatgga gacagacaca ctctgctat gggactgct gctctgggtt	60	
	ccaggtagta ctggtgacgg aagtatgcat ggagatacac ctacattgca tgaatatatg	120	
	ttagatttgc aaccagagac aactgatctc tactgttatg agcaattaaa tgacagtagt	180	
	gaggaggagg atgaaataga tggccagct ggacaagcag aaccggacag agcccattac	240	
30	aatatgttaa ccttttgttg caagtgtgac agtacgcttc ggttgtgctg acaaagtaca	300	
	cacgtagaca ttcgtacttt ggaagacctg ttaatgggca cactaggaat tgtgtgcccc	360	
	atctgcagtc agaagcccta agaattc	387	
	<210> 17		
	<211> 387		
35	<212> ADN		
	<213> Secuencia artificial		
	<220>		
	<223> Secuencia plasmídica		
	<400> 17		

ES 2 662 030 T3

```

gggtaccgccg ccaccatgga gacagacaca ctctgctat ggggtactgct gctctggggtt      60
ccaggtagca ctgggtgacgg aagcatgcat ggagatacac ctacattgca tgaatatatg      120
ttagatttgc aaccagagac aactgatctc tactgttatg agcaattaaa tgacagcagc      180
gaggaggagg atgaaataga tgggtccagct ggacaagcag aaccggacag agcccattac      240
aatattgtaa ccttttgttg caagtgtgac agcacgcttc ggttgtgcgt acaaagcaca      300
cacgtagaca ttcgtacttt ggaagacctg ttaatgggca cactaggaat tgtgtgcccc      360
atctgcagcc agaagcccta agaattc      387

```

<210> 18
 <211> 387
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Secuencia plasmídica

<400> 18

```

gggtaccgccg ccaccatgga gacagacaca ctctgctat ggggtactgct gctctggggtt      60
ccaggttcga ctgggtgacgg atcgatgcat ggagatacac ctacattgca tgaatatatg      120
ttagatttgc aaccagagac aactgatctc tactgttatg agcaattaaa tgactcgtcg      180
gaggaggagg atgaaataga tgggtccagct ggacaagcag aaccggacag agcccattac      240
aatattgtaa ccttttgttg caagtgtgac tcgacgcttc ggttgtgcgt acaatcgaca      300
cacgtagaca ttcgtacttt ggaagacctg ttaatgggca cactaggaat tgtgtgcccc      360
atctgctcgc agaagcccta agaattc      387

```

<210> 19
 <211> 387
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Secuencia plasmídica

<400> 19

```

gggtaccgccg ccaccatgga gacagacaca ctctgctat ggggtactgct gctctggggtt      60
ccaggttcaa ctgggtgacgg atcaatgcat ggagatacac ctacattgca tgaatatatg      120
ttagatttgc aaccagagac aactgatctc tactgttatg agcaattaaa tgactcatca      180
gaggaggagg atgaaataga tgggtccagct ggacaagcag aaccggacag agcccattac      240
aatattgtaa ccttttgttg caagtgtgac tcaacgcttc ggttgtgcgt acaatcaaca      300
cacgtagaca ttcgtacttt ggaagacctg ttaatgggca cactaggaat tgtgtgcccc      360
atctgctcac agaagcccta agaattc      387

```

<210> 20
 <211> 387
 <212> ADN
 <213> Secuencia artificial

<220>
<223> Secuencia plasmídica

<400> 20

5

```

ggtaccgccg ccaccatgga gacagacaca ctctgctat gggactgct gctctgggtt      60
ccagggttcta ctggtgacgg atctatgcat ggagatacac ctacattgca tgaatatatg      120

ttagatttgc aaccagagac aactgatctc tactgttatg agcaattaaa tgactcttct      180
gaggaggagg atgaaataga tgggccagct ggacaagcag aaccggacag agcccattac      240
aatattgtaa ccttttggtg caagtgtgac tctacgcttc ggttggtgctg acaatctaca      300
cacgtagaca ttcgtacttt ggaagacctg ttaatgggca cactaggaat tgtgtgcccc      360
atctgctctc agaagcccta agaattc      387

```

10

<210> 21
<211> 387
<212> ADN
<213> Secuencia artificial

15

<220>
<223> Secuencia plasmídica

<400> 21

```

ggtaccgccg ccaccatgga gacagacaca ctctgctat gggactgct gctctgggtt      60
ccagggttcca ctggtgacgg atccatgcat ggagatacac ctacattgca tgaatatatg      120
ttagatttgc aaccagagac aactgatctc tactgttatg agcaattaaa tgactcctcc      180
gaggaggagg atgaaataga tgggccagct ggacaagcag aaccggacag agcccattac      240
aatattgtaa ccttttggtg caagtgtgac tccacgcttc ggttggtgctg acaatccaca      300
cacgtagaca ttcgtacttt ggaagacctg ttaatgggca cactaggaat tgtgtgcccc      360
atctgctccc agaagcccta agaattc      387

```

20

<210> 22
<211> 387
<212> ADN
<213> Secuencia artificial

25

<220>
<223> Secuencia plasmídica

30

<400> 22

```

ggtaccgccg ccaccatgga gacggacacg ctctgctat gggactgct gctctgggtt      60
ccagggttcca cgggtgacgg atccatgcat ggagatacgc ctacgttgca tgaatatatg      120
ttagatttgc aaccagagac gacggatctc tactgttatg agcaattaaa tgacagctca      180
gaggaggagg atgaaataga tgggccagct ggacaagcag aaccggacag agcccattac      240
aatattgtaa cgttttggtg caagtgtgac tctacgcttc ggttggtgctg acaaagcacg      300
cacgtagaca ttcgtacggt ggaagacctg ttaatgggca cgctaggaat tgtgtgcccc      360
atctgctctc agaagcccta agaattc      387

```

5	<210> 23		
	<211> 387		
	<212> ADN		
	<213> Secuencia artificial		
	<220>		
10	<223> Secuencia plasmídica		
	<400> 23		
	ggtaccgccg ccaccatgga gacagacaca ctctgctat gggactgct gctctgggtt	60	
	ccagggtcca cagggtgacgg atccatgcat ggagatacac ctacattgca tgaatatatg	120	
	ttagatttgc aaccagagac aacagatctc tactgttatg agcaattaaa tgacagctca	180	
15	gaggaggagg atgaaataga tgggccagct ggacaagcag aaccggacag agcccattac	240	
	aatattgtaa cattttgttg caagtgtgac tctacacttc ggttgtgctg acaaagcaca	300	
	cacgtagaca ttcgtacatt ggaagacctg ttaatgggca cactaggaat tgtgtgcccc	360	
	atctgctctc agaagcccta agaattc	387	
	<210> 24		
20	<211> 387		
	<212> ADN		
	<213> Secuencia artificial		
	<220>		
	<223> Secuencia plasmídica		
25	<400> 24		
	ggtaccgccg ccaccatgga gactgacact ctctgctat gggactgct gctctgggtt	60	
	ccagggtcca ctggtgacgg atccatgcat ggagatactc ctactttgca tgaatatatg	120	
	ttagatttgc aaccagagac tactgatctc tactgttatg agcaattaaa tgacagctca	180	
	gaggaggagg atgaaataga tgggccagct ggacaagcag aaccggacag agcccattac	240	
30	aatattgtaa cttttgttg caagtgtgac tctactcttc ggttgtgctg acaaagcact	300	
	cacgtagaca ttcgtacttt ggaagacctg ttaatgggca ctctaggaat tgtgtgcccc	360	
	atctgctctc agaagcccta agaattc	387	
	<210> 25		
	<211> 387		
35	<212> ADN		
	<213> Secuencia artificial		
	<220>		
	<223> Secuencia plasmídica		
	<400> 25		
	ggtaccgccg ccaccatgga gaccgacacc ctctgctat gggactgct gctctgggtt	60	
	ccagggtcca cagggtgacgg atccatgcat ggagataccc ctaccttgca tgaatatatg	120	
	ttagatttgc aaccagagac caccgatctc tactgttatg agcaattaaa tgacagctca	180	

ES 2 662 030 T3

		gaggaggagg atgaaataga tgggtccagct ggacaagcag aaccggacag agcccattac	240
		aatattgtaa ccttttggtg caagtgtgac tctacccttc ggttgtgctg acaaagcacc	300
		cacgtagaca ttcgtacctt ggaagacctg ttaatgggca ccctaggaat tgtgtgcccc	360
		atctgctctc agaagcccta agaattc	387
5	<210> 26 <211> 387 <212> ADN <213> Secuencia artificial		
10	<220> <223> Secuencia plasmídica		
	<400> 26		
		ggtaccgccg ccaccatgga gacagacaca ctctgctat ggggtgctgct gctctgggtg	60
		ccagggttcca ctggtgacgg atccatgcat ggagatacac ctacattgca tgaatatatg	120
		ttagatttgc aaccagagac aactgatctc tactgttatg agcaattaaa tgacagctca	180
		gaggaggagg atgaaataga tgggtccagct ggacaagcag aaccggacag agcccattac	240
		aatattgtga ccttttggtg caagtgtgac tctacgcttc ggttgtgctg gcaaagcaca	300
		cacgtggaca ttcgtacttt ggaagacctg ttaatgggca cactaggaat tgtgtgcccc	360
		atctgctctc agaagcccta agaattc	387
15	<210> 27 <211> 387 <212> ADN <213> Secuencia artificial		
20	<220> <223> Secuencia plasmídica		
	<400> 27		
		ggtaccgccg ccaccatgga gacagacaca ctctgctat ggggtactgct gctctgggta	60
		ccagggttcca ctggtgacgg atccatgcat ggagatacac ctacattgca tgaatatatg	120
		ttagatttgc aaccagagac aactgatctc tactgttatg agcaattaaa tgacagctca	180
		gaggaggagg atgaaataga tgggtccagct ggacaagcag aaccggacag agcccattac	240
		aatattgtaa ccttttggtg caagtgtgac tctacgcttc ggttgtgctg acaaagcaca	300
		cacgtagaca ttcgtacttt ggaagacctg ttaatgggca cactaggaat tgtatgcccc	360
25		atctgctctc agaagcccta agaattc	387
30	<210> 28 <211> 387 <212> ADN <213> Secuencia artificial		
	<220> <223> Secuencia plasmídica		
35	<400> 28		

ES 2 662 030 T3

	ggtagcgccg ccaccatgga gacagacaca ctctgctat gggttctgct gctctgggtt	60
	ccagggtcca ctgggtgacgg atccatgcat ggagatacac ctacattgca tgaatatatg	120
	ttagatttgc aaccagagac aactgatctc tactgttatg agcaattaaa tgacagctca	180
	gaggaggagg atgaaataga tgggccagct ggacaagcag aaccggacag agcccattac	240
	aatattgtta ccttttggtg caagtgtgac tctacgcttc ggttggtgct tcaaagcaca	300
	cacgttgaca ttcgctacttt ggaagacctg ttaatgggca cactaggaat tgtttgcccc	360
	atctgctctc agaagcccta agaattc	387
5	<210> 29 <211> 387 <212> ADN <213> Secuencia artificial	
10	<220> <223> Secuencia plasmídica	
	<400> 29	
	ggtagcgccg ccaccatgga gacagacaca ctctgctat gggttctgct gctctgggtc	60
	ccagggtcca ctgggtgacgg atccatgcat ggagatacac ctacattgca tgaatatatg	120
	ttagatttgc aaccagagac aactgatctc tactgttatg agcaattaaa tgacagctca	180
	gaggaggagg atgaaataga tgggccagct ggacaagcag aaccggacag agcccattac	240
	aatattgtca ccttttggtg caagtgtgac tctacgcttc ggttggtgct ccaaagcaca	300
	cacgtcgaca ttcgctacttt ggaagacctg ttaatgggca cactaggaat tgtctgcccc	360
	atctgctctc agaagcccta agaattc	387
15	<210> 30 <211> 408 <212> ADN <213> Secuencia artificial	
20	<220> <223> Secuencia de enlazador plasmídico	
	<400> 30	
	ggtagcgccg ccaccatgga gaccgacacc ctctgctgt ggggtgctgct gctctgggtg	60
	cccgggtcca ccggcgacgc ggcggggcgcg ggcggggcgcg gatccatgca cggcgacacc	120
	cccacctgc acgagtacat gctggacctg cagcccgaga ccaccgacct gtactgctac	180
25	gagcagctca acgacagcag cgaggaggag gacgagatcg acggccccgc cggccaggcc	240
	gagcccgacc gcgcccacta caacatcgtg acctctgct gcaagtgcga cagcaccctg	300
	cgctctgctg tgacagagcac ccacgtggac atccgcaccc tggaggacct gctgatgggc	360
	acctggggca tcgtgtgccc catctgctcc cagaagccct aagaattc	408
30	<210> 31 <211> 408 <212> ADN <213> Secuencia artificial	

<220>
<223> Secuencia de enlazador plasmídico

<400> 31

5

```

ggtaaccgccg ccaccatgga gaccgacacc ctcttctgtgt ggggtgctgct gctctgggtg      60
cccggtctcca ccggcgacgc agcaggcgca ggcgcagcag gatccatgca cggcgacacc      120
cccaccctgc acgagtacat gctggacctg cagcccgaga ccaccgacct gtactgctac      180
gagcagctca acgacagcag cgaggaggag gacgagatcg acggccccgc cggccaggcc      240
gagcccgacc gcgcccacta caacatcgtg accttctgct gcaagtgcga cagcaccctg      300
cgctcttgcg tgcagagcac ccacgtggac atccgcaccc tggaggacct gctgatgggc      360
accctgggca tcgtgtgccc catctgctcc cagaagccct aagaattc      408

```

<210> 32
<211> 408
<212> ADN
<213> Secuencia artificial

10

<220>
<223> Secuencia de enlazador plasmídico

15

<400> 32

```

ggtaaccgccg ccaccatgga gaccgacacc ctcttctgtgt ggggtgctgct gctctgggtg      60
cccggtctcca ccggcgacgc tgcgtggcgt ggcgtgctg gatccatgca cggcgacacc      120
cccaccctgc acgagtacat gctggacctg cagcccgaga ccaccgacct gtactgctac      180
gagcagctca acgacagcag cgaggaggag gacgagatcg acggccccgc cggccaggcc      240
gagcccgacc gcgcccacta caacatcgtg accttctgct gcaagtgcga cagcaccctg      300
cgctcttgcg tgcagagcac ccacgtggac atccgcaccc tggaggacct gctgatgggc      360
accctgggca tcgtgtgccc catctgctcc cagaagccct aagaattc      408

```

<210> 33
<211> 408
<212> ADN
<213> Secuencia artificial

20

<220>
<223> Secuencia de enlazador plasmídico

25

<400> 33

```

ggtaaccgccg ccaccatgga gaccgacacc ctcttctgtgt ggggtgctgct gctctgggtg      60
cccggtctcca ccggcgacgc cgccggcgcc ggcgcgcgcg gatccatgca cggcgacacc      120
cccaccctgc acgagtacat gctggacctg cagcccgaga ccaccgacct gtactgctac      180
gagcagctca acgacagcag cgaggaggag gacgagatcg acggccccgc cggccaggcc      240
gagcccgacc gcgcccacta caacatcgtg accttctgct gcaagtgcga cagcaccctg      300
cgctcttgcg tgcagagcac ccacgtggac atccgcaccc tggaggacct gctgatgggc      360
accctgggca tcgtgtgccc catctgctcc cagaagccct aagaattc      408

```

30

	<210> 34	
	<211> 408	
	<212> ADN	
5	<213> Secuencia artificial	
	<220>	
	<223> Secuencia de enlazador plasmídico	
	<400> 34	
10	ggtagcgcgcg ccaccatgga gaccgacacc ctctgtgtgt gggtgctgct gctctgggtg	60
	cccggctcca ccggcgacag gaggggcagg ggcaggaggg gatccatgca cggcgacacc	120
	cccaccctgc acgagtacat gctggacctg cagcccagaga ccaccgacct gtactgctac	180
	gagcagctca acgacagcag cgaggaggag gacgagatcg acggccccgc cggccaggcc	240
	gagccccgacc ggcgccacta caacatcggtg accttctgct gcaagtgcga cagcacctg	300
	cgcctctgcg tgcagagcac ccacgtggac atccgcaccc tggaggacct gctgatgggc	360
	accctgggca tcgtgtgccc catctgctcc cagaagccct aagaattc	408
	<210> 35	
	<211> 408	
15	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Secuencia de enlazador plasmídico	
20	<400> 35	
	ggtagcgcgcg ccaccatgga gaccgacacc ctctgtgtgt gggtgctgct gctctgggtg	60
	cccggctcca ccggcgacag aagaggcaga ggcagaagag gatccatgca cggcgacacc	120
	cccaccctgc acgagtacat gctggacctg cagcccagaga ccaccgacct gtactgctac	180
	gagcagctca acgacagcag cgaggaggag gacgagatcg acggccccgc cggccaggcc	240
	gagccccgacc ggcgccacta caacatcggtg accttctgct gcaagtgcga cagcacctg	300
	cgcctctgcg tgcagagcac ccacgtggac atccgcaccc tggaggacct gctgatgggc	360
25	accctgggca tcgtgtgccc catctgctcc cagaagccct aagaattc	408
	<210> 36	
	<211> 408	
30	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Secuencia de enlazador plasmídico	
35	<400> 36	

ES 2 662 030 T3

	ggtagcgccg ccaccatgga gaccgacacc ctctgtgtgt ggggtgctgct gctctgggtg	60
	cccggctcca ccggcgaccg gcggggccgg ggccggcggg gatccatgca cggcgacacc	120
	cccaccctgc acgagtacat gctggacctg cagcccgaga ccaccgacct gtactgttac	180
	gagcagctca acgacagcag cgaggaggag gacgagatcg acggccccgc cggccaggcc	240
	gagcccgacc gcgcccacta caacatcgtg accttctgct gcaagtgcga cagcaccctg	300
	cgctctgctg tgcagagcac ccacgtggac atccgcaccc tggaggacct gctgatgggc	360
	accctgggca tcgtgtgccc catctgtctcc cagaagccct aagaattc	408
5	<210> 37 <211> 408 <212> ADN <213> Secuencia artificial	
10	<220> <223> Secuencia de enlazador plasmídico	
	<400> 37	
	ggtagcgccg ccaccatgga gaccgacacc ctctgtgtgt ggggtgctgct gctctgggtg	60
	cccggctcca ccggcgaccg acgaggccga ggccgacgag gatccatgca cggcgacacc	120
	cccaccctgc acgagtacat gctggacctg cagcccgaga ccaccgacct gtactgttac	180
	gagcagctca acgacagcag cgaggaggag gacgagatcg acggccccgc cggccaggcc	240
	gagcccgacc gcgcccacta caacatcgtg accttctgct gcaagtgcga cagcaccctg	300
	cgctctgctg tgcagagcac ccacgtggac atccgcaccc tggaggacct gctgatgggc	360
	accctgggca tcgtgtgccc catctgtctcc cagaagccct aagaattc	408
15	<210> 38 <211> 408 <212> ADN <213> Secuencia artificial	
20	<220> <223> Secuencia de enlazador plasmídico	
	<400> 38	
	ggtagcgccg ccaccatgga gaccgacacc ctctgtgtgt ggggtgctgct gctctgggtg	60
	cccggctcca ccggcgaccg tcgtggccgt ggccgtcgtg gatccatgca cggcgacacc	120
	cccaccctgc acgagtacat gctggacctg cagcccgaga ccaccgacct gtactgttac	180
	gagcagctca acgacagcag cgaggaggag gacgagatcg acggccccgc cggccaggcc	240
	gagcccgacc gcgcccacta caacatcgtg accttctgct gcaagtgcga cagcaccctg	300
	cgctctgctg tgcagagcac ccacgtggac atccgcaccc tggaggacct gctgatgggc	360
25	accctgggca tcgtgtgccc catctgtctcc cagaagccct aagaattc	408
30	<210> 39 <211> 408 <212> ADN <213> Secuencia artificial	
	<220> <223> Secuencia de enlazador plasmídico	

<400> 39

```

    ggtaccgccg ccaccatgga gaccgacacc ctctgctgt ggggtgctgct gctctgggtg      60
    cccggctcca ccggcgaccg ccggcgccgc ggccgccgc gatccatgca cggcgacacc      120
    cccaccctgc acgagtacat gctggacctg cagcccgaga ccaccgacct gtactgctac      180
    gagcagctca acgacagcag cgaggaggag gacgagatcg acggccccgc cggccaggcc      240
    gagcccgacc gcgcccacta caacatcgtg accttctgct gcaagtgcga cagcaccctg      300
    cgcctctgcg tgcagagcac ccacgtggac atccgcaccc tggaggacct gctgatgggc      360
    accctgggca tcgtgtgccc catctgctcc cagaagccct aagaattc      408
  
```

5 <210> 40
 <211> 408
 <212> ADN
 <213> Secuencia artificial

10 <220>
 <223> Secuencia de enlazador plasmídico

<400> 40

```

    ggtaccgccg ccaccatgga gaccgacacc ctctgctgt ggggtgctgct gctctgggtg      60
    cccggctcca ccggcgacaa taatggcaat ggcaataatg gatccatgca cggcgacacc      120
    cccaccctgc acgagtacat gctggacctg cagcccgaga ccaccgacct gtactgctac      180
    gagcagctca acgacagcag cgaggaggag gacgagatcg acggccccgc cggccaggcc      240
    gagcccgacc gcgcccacta caacatcgtg accttctgct gcaagtgcga cagcaccctg      300
    cgcctctgcg tgcagagcac ccacgtggac atccgcaccc tggaggacct gctgatgggc      360
    accctgggca tcgtgtgccc catctgctcc cagaagccct aagaattc      408
  
```

20 <210> 41
 <211> 408
 <212> ADN
 <213> Secuencia artificial

25 <220>
 <223> Secuencia de enlazador plasmídico

<400> 41

```

    ggtaccgccg ccaccatgga gaccgacacc ctctgctgt ggggtgctgct gctctgggtg      60
    cccggctcca ccggcgacaa caacggcaac ggcaacaacg gatccatgca cggcgacacc      120
    cccaccctgc acgagtacat gctggacctg cagcccgaga ccaccgacct gtactgctac      180
    gagcagctca acgacagcag cgaggaggag gacgagatcg acggccccgc cggccaggcc      240
    gagcccgacc gcgcccacta caacatcgtg accttctgct gcaagtgcga cagcaccctg      300
    cgcctctgcg tgcagagcac ccacgtggac atccgcaccc tggaggacct gctgatgggc      360
    accctgggca tcgtgtgccc catctgctcc cagaagccct aagaattc      408
  
```

30 <210> 42
 <211> 408
 <212> ADN
 <213> Secuencia artificial

35

ES 2 662 030 T3

<220>

<223> Secuencia de enlazador plasmídico

<400> 42

5

```

ggtagcgccg ccaccatgga gaccgacacc ctctgctgt gggtagctgt gctctgggtg      60
ccgggtcca ccggcgacca tcatggccat ggccatcatg gatccatgca cggcgacacc      120
cccaccctgc acgagtacat gctggacctg cagcccgaga ccaccgacct gtactgttac      180
gagcagctca acgacagcag cgaggaggag gacgagatcg acggccccgc cggccaggcc      240
gagcccgacc gcgcccacta caacatcgtg accttctgt gcaagtgcga cagcaccctg      300
cgctctgctg tgcagagcac ccacgtggac atccgcaccc tggaggacct gctgatgggc      360
accctgggca tcgtgtgccc catctgctcc cagaagccct aagaattc      408

```

<210> 43

<211> 408

10

<212> ADN

<213> Secuencia artificial

<220>

<223> Secuencia de enlazador plasmídico

15

<400> 43

```

ggtagcgccg ccaccatgga gaccgacacc ctctgctgt gggtagctgt gctctgggtg      60
ccgggtcca ccggcgacca ccacggccac ggccaccacg gatccatgca cggcgacacc      120
cccaccctgc acgagtacat gctggacctg cagcccgaga ccaccgacct gtactgttac      180
gagcagctca acgacagcag cgaggaggag gacgagatcg acggccccgc cggccaggcc      240
gagcccgacc gcgcccacta caacatcgtg accttctgt gcaagtgcga cagcaccctg      300
cgctctgctg tgcagagcac ccacgtggac atccgcaccc tggaggacct gctgatgggc      360
accctgggca tcgtgtgccc catctgctcc cagaagccct aagaattc      408

```

20

<210> 44

<211> 408

<212> ADN

<213> Secuencia artificial

25

<220>

<223> Secuencia de enlazador plasmídico

<400> 44

30

```

ggtagcgccg ccaccatgga gaccgacacc ctctgctgt gggtagctgt gctctgggtg      60
ccgggtcca ccggcgacaa gaagggcaag ggcaagaagg gatccatgca cggcgacacc      120
cccaccctgc acgagtacat gctggacctg cagcccgaga ccaccgacct gtactgttac      180
gagcagctca acgacagcag cgaggaggag gacgagatcg acggccccgc cggccaggcc      240
gagcccgacc gcgcccacta caacatcgtg accttctgt gcaagtgcga cagcaccctg      300
cgctctgctg tgcagagcac ccacgtggac atccgcaccc tggaggacct gctgatgggc      360
accctgggca tcgtgtgccc catctgctcc cagaagccct aagaattc      408

```

5	<210> 45		
	<211> 408		
	<212> ADN		
	<213> Secuencia artificial		
10	<220>		
	<223> Secuencia de enlazador plasmídico		
	<400> 45		
	ggtagcgcgc ccaccatgga gaccgacacc ctctgtgtgt gggtagctgt gctctgggtg	60	
15	cccggtcca ccggcgacaa aaaaggcaaa ggcaaaaag gatccatgca cggcgacacc	120	
	cccacctgc acgagtacat gctggacctg cagcccgaga ccaccgacct gtactgtac	180	
	gagcagctca acgacagcag cgaggaggag gacgagatcg acggccccgc cggccaggcc	240	
	gagcccgacc gcgcccacta caacatcgtg accttctgt gcaagtgcga cagcaccctg	300	
20	cgcctctgcg tgcagagcac ccacgtggac atccgcaccc tggaggacct gctgatgggc	360	
	accctgggca tcgtgtgcc catctgtctc cagaagccct aagaattc	408	
	<210> 46		
	<211> 408		
	<212> ADN		
	<213> Secuencia artificial		
25	<220>		
	<223> Secuencia de enlazador plasmídico		
	<400> 46		
	ggtagcgcgc ccaccatgga gaccgacacc ctctgtgtgt gggtagctgt gctctgggtg	60	
30	cccggtcca ccggcgactt ttttggcttt ggcttttttg gatccatgca cggcgacacc	120	
	cccacctgc acgagtacat gctggacctg cagcccgaga ccaccgacct gtactgtac	180	
	gagcagctca acgacagcag cgaggaggag gacgagatcg acggccccgc cggccaggcc	240	
	gagcccgacc gcgcccacta caacatcgtg accttctgt gcaagtgcga cagcaccctg	300	
35	cgcctctgcg tgcagagcac ccacgtggac atccgcaccc tggaggacct gctgatgggc	360	
	accctgggca tcgtgtgcc catctgtctc cagaagccct aagaattc	408	
	<210> 47		
	<211> 408		
	<212> ADN		
	<213> Secuencia artificial		
40	<220>		
	<223> Secuencia de enlazador plasmídico		
	<400> 47		

ES 2 662 030 T3

	ggtagcgccg ccaccatgga gaccgacacc ctctgctgt gggtagctgt gctctgggtg	60
	cccggtctcca ccggcgactt ctccggttc ggcttcttcg gatccatgca cggcgacacc	120
	cccacctgc acgagtacat gctggacctg cagcccgaga ccaccgacct gtactgttac	180
	gagcagctca acgacagcag cgaggaggag gacgagatcg acggccccgc cggccaggcc	240
	gagcccgacc gcgcccacta caacatcgtg accttctgct gcaagtgcga cagcacctg	300
	cgctctgcg tgcagagcac ccacgtggac atccgcaccc tggaggacct gctgatgggc	360
	acctgggca tcgtgtgccc catctgctcc cagaagccct aagaattc	408
5	<210> 48 <211> 408 <212> ADN <213> Secuencia artificial	
10	<220> <223> Secuencia de enlazador plasmídico	
	<400> 48	
	ggtagcgccg ccaccatgga gaccgacacc ctctgctgt gggtagctgt gctctgggtg	60
	cccggtctcca ccggcgacta ttatggctat ggctattatg gatccatgca cggcgacacc	120
	cccacctgc acgagtacat gctggacctg cagcccgaga ccaccgacct gtactgttac	180
	gagcagctca acgacagcag cgaggaggag gacgagatcg acggccccgc cggccaggcc	240
	gagcccgacc gcgcccacta caacatcgtg accttctgct gcaagtgcga cagcacctg	300
	cgctctgcg tgcagagcac ccacgtggac atccgcaccc tggaggacct gctgatgggc	360
15	acctgggca tcgtgtgccc catctgctcc cagaagccct aagaattc	408
20	<210> 49 <211> 408 <212> ADN <213> Secuencia artificial	
	<220> <223> Secuencia de enlazador plasmídico	
25	<400> 49	
	ggtagcgccg ccaccatgga gaccgacacc ctctgctgt gggtagctgt gctctgggtg	60
	cccggtctcca ccggcgacta ctacggctac ggctactacg gatccatgca cggcgacacc	120
	cccacctgc acgagtacat gctggacctg cagcccgaga ccaccgacct gtactgttac	180
	gagcagctca acgacagcag cgaggaggag gacgagatcg acggccccgc cggccaggcc	240
	gagcccgacc gcgcccacta caacatcgtg accttctgct gcaagtgcga cagcacctg	300
	cgctctgcg tgcagagcac ccacgtggac atccgcaccc tggaggacct gctgatgggc	360
	acctgggca tcgtgtgccc catctgctcc cagaagccct aagaattc	408
30	<210> 50 <211> 1707 <212> ADN <213> Virus	

<400> 50

atggagaaaa tagtgcttct ttttgcaata gtcagtcttg ttaaaagtga tcagatttgc	60
attggttacc atgcaaaca ctcgacagag cagggttgaca caataatgga aaagaacggt	120
actgttacac atgccccaga catactggaa aagacacaca acggaagct ctgcgatcta	180
gatggagtga agcctcta atttgagagat tgtagtgtag ctggatggct cctcggaac	240
ccaatgtgtg acgaattcat caatgtgccg gaatggtctt acatagtga gaaggccaat	300
ccagccaatg acctctgtta ccaggggat ttcaacgact atgaagaatt gaaacaccta	360
ttgagcagaa taaaccattt tgagaaaatt cagatcatcc caaaaattc ttggtccagt	420
catgaagcct cattaggggt gagctcagca tgtccatacc aaggaaagtc ctctttttc	480
aggaatgtgg tatggcttat caaaaagaac aatgcatacc caacaataaa gaggagctac	540
aataatacca accaagaaga tcttttggtt ttgtggggga ttcacatcc taatgatgcg	600
gcagagcaga ctaggctcta tcaaaaccca accacctaca tttcogttgg gacatcaaca	660
ctaaaccaga gattggtacc aaaaatagct actagatcca aagtaaagcg gcaaaatgga	720
aggatggagt tcttctggac aattttaaaa ccgaatgatg caatcaactt cgagagcaat	780
ggaaatttca ttgctccaga atatgcatac aaaattgtca agaaagggga ctcagcaatt	840
atgaaaagtg aattggaata tggtaactgc aacaccaagt gtcaaaactc aatggggcg	900
ataaactcta gtatgccatt ccacaatata caccctctca ccatcgggga atgccccaa	960
tatgtgaaat caaacagatt agtccttgcg actgggtcca gaaatagccc tcaaagagag	1020
agaagaagaa aaaagagagg attatttggg gctatagcag gttttataga gggaggatgg	1080
cagggaatgg tagatggtg gtatgggtac caccatagca atgagcaggg gagtgggtac	1140
gctgcagaca aagaatccac tcaaaaggca atagatggag tcaccaataa ggtcaactcg	1200
atcattgaca aaatgaacac tcagtttgag gccgttgga gggaatttaa taacttagaa	1260
aggagaatag agaattttaa caagaagatg gaagacggat tcctagatgt ctggacttat	1320
aatgtgaac ttctggttct catggaaaat gagagaactc tagactttca tgactcaaat	1380
gtcaagaacc ttacgacaa ggtccgacta cagcttaggg ataattgcaa ggagctgggt	1440
aacggttgtt tcgagttcta tcacaaatgt gataatgaat gtatggaaag tgtaagaaac	1500
ggaacgtatg actacccgca gtattcagaa gaagcaagac taaaagaga ggaaataagt	1560
ggagtaaaat tggagtcaat aggaacttac caaatactgt caattttattc tacagtggcg	1620
agttccctag cactggcaat catggtagct ggtctatctt tatggatgtg ctccaatggg	1680
tcgttacaat gcagaatttg catttaa	1707

5

<210> 51

<211> 568

<212> PRT

<213> Virus

10

<400> 51

ES 2 662 030 T3

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

ES 2 662 030 T3

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 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
 145 150 155 160

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
 225 230 235 240

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 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

ES 2 662 030 T3

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 Arg Arg Arg Lys Lys Arg Gly Leu Phe Gly Ala Ile
 340 345 350
 Ala Gly Phe Ile Glu Gly Gly Trp Gln Gly Met Val Asp Gly Trp Tyr
 355 360 365
 Gly Tyr His His Ser Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Lys
 370 375 380
 Glu Ser Thr Gln Lys Ala Ile Asp Gly Val Thr Asn Lys Val Asn Ser
 385 390 395 400
 Ile Ile Asp Lys Met Asn Thr Gln Phe Glu Ala Val Gly Arg Glu Phe
 405 410 415
 Asn Asn Leu Glu Arg Arg Ile Glu Asn Leu Asn Lys Lys Met Glu Asp
 420 425 430
 Gly Phe Leu Asp Val Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Met
 435 440 445
 Glu Asn Glu Arg Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu
 450 455 460
 Tyr Asp Lys Val Arg Leu Gln Leu Arg Asp Asn Ala Lys Glu Leu Gly
 465 470 475 480
 Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp Asn Glu Cys Met Glu
 485 490 495
 Ser Val Arg Asn Gly Thr Tyr Asp Tyr Pro Gln Tyr Ser Glu Glu Ala
 500 505 510
 Arg Leu Lys Arg Glu Glu Ile Ser Gly Val Lys Leu Glu Ser Ile Gly
 515 520 525
 Thr Tyr Gln Ile Leu Ser Ile Tyr Ser Thr Val Ala Ser Ser Leu Ala
 530 535 540
 Leu Ala Ile Met Val Ala Gly Leu Ser Leu Trp Met Cys Ser Asn Gly
 545 550 555 560
 Ser Leu Gln Cys Arg Ile Cys Ile
 565

<210> 52
<211> 1707
<212> ADN
<213> Virus

5

<400> 52

atggaaaaaa tcgtgctgct gttecgctatc gtctcgctgg tcaaatcgga tcagatctgc	60
atcggatacc atgctaacaa ctcgacggaa caggctcgaca cgatcatgga aaagaacgtc	120
acggtcacgc atgtctcaaga catcctggaa aagacgcaca acggaaagct gtgcgatctg	180
gatggagtga agcctctgat cctgagagat tgttcggctg ctggatggct gctgggaaac	240
cccatgtgtg acgaatttat caatgtgccc gaatggctgt acatcgtgga aaaggctaata	300
cccgcataatg acctgtgtta ccccgagat tttaacgact atgaagaact gaaacacctg	360
ctgtcgagaa tcaaccattt cgaaaaaatc cagatcatcc ccaaaaattc gtggctgctg	420
catgaagctt cgctgggagt gtctcggtt tgtccctacc aaggaaagtc gtcgttcttt	480
agaaatgtgg tctggctgat caaaaagaac aatgcttacc ccacgatcaa gagatcgtac	540
aataatacga accaagaaga tctgctggtc ctgtggggaa tccaccatcc taatgatgct	600
gctgaacaga cgagactgta tcaaaacccc acgacgtaca tctcggtcgg aacgtcgacg	660
ctgaaccaga gactgggtccc caaaatcgct acgagatcga aagtcaacgg acaaaatgga	720
agaatggaat ttttttggac gatcctgaaa cccaatgatg ctatcaactt tgaatcgaat	780
ggaaatttta tcgctcccga atatgcttac aaaatcgta agaaaggaga ctcggtctatc	840
atgaaatcgg aactggaata tggaaactgc aacacgaagt gtcaaacgcc catgggagct	900
atcaactcgt cgatgccctt tcacaatatc caccctctga cgatcggaga atgccccaaa	960
tatgtgaaat cgaacagact ggtcctggct acgggactga gaaattcgcc tcaaagagaa	1020
agaagaagaa aaaagagagg actgttcgga gctatcgctg gattcatcga aggaggatgg	1080
cagggaatgg tcgatggatg gtatggatac caccattcga atgaacaggg atcgggatac	1140
gctgctgaca aagaatcgac gcaaaaggct atcgatggag tcacgaataa ggtcaactcg	1200
atcatcgaca aaatgaacac gcagttcgaa gctgtcgga gagaattcaa taacctggaa	1260
agaagaatcg aaaatctgaa caagaagatg gaagacggat ttctggatgt ctggacgtat	1320
aatgctgaac tgctggctct gatggaaaat gaaagaacgc tggacttcca tgactcgaat	1380
gtcaagaacc tgtacgacaa ggtccgactg cagctgagag ataatgctaa ggaactggga	1440
aacggatggt ttgaatttta tcacaaatgt gataatgaat gtatggaatc ggtcagaaac	1500
ggaacgtatg actaccccca gtattcggaa gaagctagac tgaaaagaga agaaatctcg	1560
ggagtcaaac tggaatcgat cggaacgtac caaatcctgt cgatctattc gacggtggct	1620
tcgtcgctgg ctctggctat catggctcgt ggactgtcgc tgtggatgtg ctggaatgga	1680
tcgtcgcaat gcagaatctg catctaa	1707

10

<210> 53
<211> 1701
<212> ADN
<213> Virus

15

<400> 53

```

atgaagacta tcattgctct gagctacatt ttatgtctgg tcttcgctca aaaacttccc      60
cgaaatgaca acagcacggc aacgctgtgc ttgggacacc atgcagtgtc aaacggaaca      120
ctagtgaaaa caatcacgaa tgaccaaatt gaagtgacta atgctactga attggttcag      180
agttcctcaa caggtagaat atgtgaccga cctcatcgaa tccttgatgg ggaaaactgc      240
acactgatag atgctctctt gggagaccct cattgtgata gtttccaaaa caaggaatgg      300
gacctttttg tagaacgcag cacagcttac agcgactgtt acccttatga tgtgccggat      360
tatgcctccc ttaggtcact agttgcctca tccggcacc tggagttaa cgatgaaagt      420
ttcgattgga ctggagtctc tcaggatgga acaagcaatg cttgcaaaag gagatctgtt      480
aaaagttttt ttagtagatt aaattggttg tacaaattag aatacaataa tccagcactg      540
aacgtgacta tgccaaacaa tgaaaaattt gacaaattgt acatttgggg ggtgcaccac      600
ccgagcacgg acagtgacca aaccagtcta tatgttcaag catcaggagg agtcacaatc      660
tctacaaaaa gaagccaaca aactgtaatc ccgaatatcg gatctagacc ctgggtaagg      720
ggtatctcca gcagaataag catctattgg acaatagtaa aacctggaga catacttatg      780
attaacagca cagggaatct aatcgccctt cggggttact tcaagatacg aagtggagaa      840
agctcaataa tgaggtcaga tgcaccatt gatagctgca attctgaatg catcactcca      900
aatggaagca ttcccaataa caaaccattt caaaatgtaa acaggatcac atatggggcc      960
tgtcctagat atgttaaaca aaaaactcta aaattggcaa cagggatgcy gaatgtacca     1020
gagaaacaag ctaggggcat attcggcgcc atcgcagggt tcatagaaaa tggttgggag     1080
ggaatggtag acggttggtg cggttttagg catctaaatt ctgagggctc aggacaagca     1140
gcagacctca aaagcactca ggcagcaatt aaccaaatac acgggaaact gaataggttg     1200
gtcgaaaaaa caaacgagaa attocatcaa attgaaaaag aattctcaga cgtggaaggg     1260
agaattcagg atctcgagaa atatgttgaa gacacaaaa tagatctctg gtcatacaat     1320
gcggagcttc ttgttgccct ggagaaccaa cacacaattg atctaactga ctcagaaatg     1380
aacaactgt tcgaaagaac aaggaaacaa ctgagggaaa atgctgagga catgggcaat     1440
ggttgcttca aaatatacca caaatgtgac aatgcctgca tagggtcgat cagaaatgga     1500
acttatgacc ataatgtata cagagacgaa gcattaaaca accgactcca tatcaaaggg     1560
gttgagctga agtcaggata caaagattgg atcttatgga tctcattttc catatcatgc     1620
tttttgtttt gtgttgtttt gctggggttc atcatgtggg cctgccaaaa aggcaacatt     1680
aggtgcaaca tttgcatttg a                                     1701

```

5

<210> 54

<211> 566

<212> PRT

<213> Virus

10

<400> 54

ES 2 662 030 T3

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

ES 2 662 030 T3

Lys Ser Phe Phe Ser Arg Leu Asn Trp Leu Tyr Lys Leu Glu Tyr Lys
 165 170 175
 Tyr Pro Ala Leu Asn Val Thr Met Pro Asn Asn Glu Lys Phe Asp Lys
 180 185 190
 Leu Tyr Ile Trp Gly Val His His Pro Ser Thr Asp Ser Asp Gln Thr
 195 200 205
 Ser Leu Tyr Val Gln Ala Ser Gly Arg Val Thr Ile Ser Thr Lys Arg
 210 215 220
 Ser Gln Gln Thr Val Ile Pro Asn Ile Gly Ser Arg Pro Trp Val Arg
 225 230 235 240
 Gly Ile Ser Ser Arg Ile Ser Ile Tyr Trp Thr Ile Val Lys Pro Gly
 245 250 255
 Asp Ile Leu Met Ile Asn Ser Thr Gly Asn Leu Ile Ala Pro Arg Gly
 260 265 270
 Tyr Phe Lys Ile Arg Ser Gly Glu Ser Ser Ile Met Arg Ser Asp Ala
 275 280 285
 Pro Ile Asp Ser Cys Asn Ser Glu Cys Ile Thr Pro Asn Gly Ser Ile
 290 295 300
 Pro Asn Asn Lys Pro Phe Gln Asn Val Asn Arg Ile Thr Tyr Gly Ala
 305 310 315 320
 Cys Pro Arg Tyr Val Lys Gln Lys Thr Leu Lys Leu Ala Thr Gly Met
 325 330 335
 Arg Asn Val Pro Glu Lys Gln Ala Arg Gly Ile Phe Gly Ala Ile Ala
 340 345 350
 Gly Phe Ile Glu Asn Gly Trp Glu Gly Met Val Asp Gly Trp Tyr Gly
 355 360 365
 Phe Arg His Leu Asn Ser Glu Gly Ser Gly Gln Ala Ala Asp Leu Lys
 370 375 380
 Ser Thr Gln Ala Ala Ile Asn Gln Ile Asn Gly Lys Leu Asn Arg Leu
 385 390 395 400

ES 2 662 030 T3

Val Glu Lys Thr Asn Glu Lys Phe His Gln Ile Glu Lys Glu Phe Ser
405 410 415

Asp Val Glu Gly Arg Ile Gln Asp Leu Glu Lys Tyr Val Glu Asp Thr
420 425 430

Lys Ile Asp Leu Trp Ser Tyr Asn Ala Glu Leu Leu Val Ala Leu Glu
435 440 445

Asn Gln His Thr Ile Asp Leu Thr Asp Ser Glu Met Asn Lys Leu Phe
450 455 460

Glu Arg Thr Arg Lys Gln Leu Arg Glu Asn Ala Glu Asp Met Gly Asn
465 470 475 480

Gly Cys Phe Lys Ile Tyr His Lys Cys Asp Asn Ala Cys Ile Gly Ser
485 490 495

Ile Arg Asn Gly Thr Tyr Asp His Asn Val Tyr Arg Asp Glu Ala Leu
500 505 510

Asn Asn Arg Leu His Ile Lys Gly Val Glu Leu Lys Ser Gly Tyr Lys
515 520 525

Asp Trp Ile Leu Trp Ile Ser Phe Ser Ile Ser Cys Phe Leu Phe Cys
530 535 540

Val Val Leu Leu Gly Phe Ile Met Trp Ala Cys Gln Lys Gly Asn Ile
545 550 555 560

Arg Cys Asn Ile Cys Ile
565

<210> 55
<211> 1701
<212> ADN
<213> Virus

<400> 55

atgaagacga tcatcgctct gtcgtacatc ctgtgtctgg tctttgctca aaaactgccc	60
cgaaatgaca actcgacggc tacgctgtgc ctgggacacc atgctgtgtc gaacggaacg	120
ctggtgaaaa cgatcacgaa tgaccaaata gaagtgcga atgctacgga actggtccag	180
tcgtcgtcga cgggaagaat ctgtgaccga cctcatcgaa tcctggatgg agaaaactgc	240
acgctgatcg atgctctgct gggagaccct cattgtgatt cgtttcaaaa caaggaatgg	300

gacctgttcg	tcgaacgctc	gacggcttac	tcggactggt	acccttatga	tgtgcccgat	360
tatgcttcgc	tgagatcgct	ggtcgcttcg	tcgggaacgc	tggaattcaa	cgatgaatcg	420
tttgattgga	cgggagtctc	gcaggatgga	acgtcgaatg	cttgcaaaag	aagatcggtc	480
aaatcgttct	tctcgagact	gaattggctg	tacaaactgg	aatacaata	tcccgcctcg	540
aacgtgacga	tgcccaacaa	tgaaaaattc	gacaaactgt	acatctgggg	agtgcaccac	600
ccctcgacgg	actcggacca	aacgtcgctg	tatgtccaag	cttcgggaag	agtcacgac	660
tcgacgaaaa	gatcgcaaca	aacggtcatc	cccaatatcg	gatcgagacc	ctgggtcaga	720
ggaatctcgt	cgagaatctc	gatctattgg	acgatcgtca	aacctggaga	catcctgatg	780
atcaactcga	cgggaaatct	gatcgctcct	cgaggatact	ttaagatccg	atcggggaga	840
tcgtcgatca	tgagatcgga	tgtcccatc	gattcgtgca	attcggaatg	catcacgccc	900
aatggatcga	tccccataa	caaacccttc	caaatgtca	acagaatcac	gtatggagct	960
tgtcctagat	atgtcaaaca	aaaaacgctg	aaactggcta	cgggaatgcg	aatgtcccc	1020
gaaaaacaag	ctagaggaat	ctttggagct	atcgctggat	ttatcgaaaa	tggatgggaa	1080
ggaatggctg	acggatggta	cggattcaga	catctgaatt	cgggaaggatc	gggacaagct	1140
gctgacctga	aatcgacgca	ggctgctatc	aaccaaata	acggaaaact	gaatagactg	1200
gtcgaaaaaa	cgaacgaaaa	atttcatcaa	atcgaaaaag	aattttcgga	cgtggaagga	1260
agaatccagg	atctggaaaa	atatgtcgaa	gacacgaaaa	tcgatctgtg	gtcgtacaat	1320
gctgaactgc	tggtcgctct	ggaaaaccaa	cacacgatcg	atctgacgga	ctcggaaatg	1380
aacaaactgt	ttgaaagaac	gagaaaacaa	ctgagagaaa	atgctgaaga	catgggaaat	1440
ggatgcttta	aaatctacca	caaatgtgac	aatgcttgca	tcggatcgat	cagaaatgga	1500
acgtatgacc	ataatgtcta	cagagacgaa	gctctgaaca	accgactgca	tatcaaagga	1560
gtcgaactga	agtcgggata	caaagattgg	atcctgtgga	tctcgttctc	gatctcgtgc	1620
ttcctgttct	gtgtcgtcct	gctgggattt	atcatgtggg	cttgccaaaa	aggaaacatc	1680
agatgcaaca	tctgcactcg	a				1701

<210> 56
 <211> 1410
 <212> ADN
 <213> Virus

<400> 56

atgaatccaa	atcagaagat	aacaaccatt	ggatcaatct	gtatggtaat	tggaatagtt	60
agcttgatgt	tacaaattgg	gaacataatc	tcaatatggg	ttagtcattc	aattcaaaca	120
gggaatcaac	accaggctga	accatgcaat	caaagcatta	ttacttatga	aaacaacacc	180

ES 2 662 030 T3

```

tgggtaaacc agacatatgt caacatcagc aataccaatt ttcttactga gaaagctgtg      240
gcttcagtaa cattagcggg caattcatct ctttgcccca ttagtggatg ggctgtatac      300
agtaaggaca acggtataag aatcgggtcc aaggggggatg tgtttgttat aagagagccg      360
ttcatctcat gctccactt ggaatgcaga actttctttt tgactcaggg agccttgctg      420
aatgacaagc attctaattg gaccgtcaaa gacagaagcc ctcacagaac attaatgagt      480
tgtcccgtag gtgaggctcc ttccccatag aactcgaggt ttgagtctgt tgcttggtcg      540
gcaagtgctt gtcagtatgg cactagttag ttgacaattg gaatttctgg ccagagacaat      600
ggggctgtgg ctgtattgaa atacaatggc ataataacag acactatcaa gagttggagg      660
aacaacataa tgagaactca agagtctgaa tgtgcatgtg taaatggctc ttgctttact      720
gttatgactg atggaccaag taatgggcag gcttcataca aaatcttcag aatagaaaaa      780
gggaaagtag ttaaatcagc cgaattaaat gcccctaatt atcactatga ggagtgtcc      840
tgttatcctg atgctggaga aatcacatgt gtgtgcaggg ataactggca tggctcaaat      900
cggccatggg tatctttcaa tcaaaatttg gagtatcgaa taggatatat atgcagtgga      960
gttttcggag acaatccacg cccaatgat gggacaggca gttgtggtcc ggtgtcccct     1020
aaaggggcat atggaataaa agggttctca tttaaatacg gcaatggtgt ttggatcggg     1080
agaacaaaaa gcactaatc caggagcggc tttgaaatga tttgggatcc aaatggatgg     1140
actggtacgg acagtaattt ttcagtaaag caagatattg tagctataac cgattggtca     1200
ggatatagcg ggagttttgt ccagcatcca gaactgacag gattagattg cataagacct     1260
tgtttctggg ttgagctaag cagagggcgg ccaaagaga gcacaatttg gactagtggg     1320
agcagcatat ccttttgtgg tgtaaatagt gacactgtgg gttggtcttg gccagacggt     1380
gctgagttgc cattcaccat tgacaagtag                                     1410

```

<210> 57
 <211> 469
 <212> PRT
 <213> Virus

<400> 57

```

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

```

ES 2 662 030 T3

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	195	200	205	
Asn	Gly	Ile	Ile	Thr	Asp	Thr	Ile	Lys	Ser	Trp	Arg	Asn	Asn	Ile	Met	210	215	220	
Arg	Thr	Gln	Glu	Ser	Glu	Cys	Ala	Cys	Val	Asn	Gly	Ser	Cys	Phe	Thr	225	230	235	240
Val	Met	Thr	Asp	Gly	Pro	Ser	Asn	Gly	Gln	Ala	Ser	Tyr	Lys	Ile	Phe	245	250	255	
Arg	Ile	Glu	Lys	Gly	Lys	Val	Val	Lys	Ser	Ala	Glu	Leu	Asn	Ala	Pro	260	265	270	
Asn	Tyr	His	Tyr	Glu	Glu	Cys	Ser	Cys	Tyr	Pro	Asp	Ala	Gly	Glu	Ile	275	280	285	

Thr Cys Val Cys Arg Asp Asn Trp His Gly Ser Asn Arg Pro Trp Val
 290 295 300

Ser Phe Asn Gln Asn Leu Glu Tyr Arg Ile Gly Tyr Ile Cys Ser Gly
 305 310 315 320

Val Phe Gly Asp Asn Pro Arg Pro Asn Asp Gly Thr Gly Ser Cys Gly
 325 330 335

Pro Val Ser Pro Lys Gly Ala Tyr Gly Ile Lys Gly Phe Ser Phe Lys
 340 345 350

Tyr Gly Asn Gly Val Trp Ile Gly Arg Thr Lys Ser Thr Asn Ser Arg
 355 360 365

Ser Gly Phe Glu Met Ile Trp Asp Pro Asn Gly Trp Thr Gly Thr Asp
 370 375 380

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

Gly Tyr Ser Gly Ser Phe Val Gln His Pro Glu Leu Thr Gly Leu Asp
 405 410 415

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

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

Asn Ser Asp Thr Val Gly Trp Ser Trp Pro Asp Gly Ala Glu Leu Pro
 450 455 460

Phe Thr Ile Asp Lys
 465

<210> 58
 <211> 1410
 <212> ADN
 <213> Virus

<400> 58

atgaatccca atcagaagat caccgacgatc ggatcgatct gtatgggtcat cggaatcgtc 60
 tcgctgatgc tgcaaactcg aaacatcatc tcgatctggg tctcgcatte gatccaaacg 120
 ggaaatcaac accaggctga accctgcaat caatcgatca tcacgtatga aaacaacacg 180
 tgggtcaacc agacgtatgt caacatctcg aatacgaatt tcctgacgga aaaagctgtg 240

gcttcggtca cgctggctgg aaattcgctg ctgtgcccc tctcgggatg ggctgtctac 300
tcgaaggaca acggaatcag aatcggatcg aaggagatg tgttcgtcat cagagaaccc 360
tttatctcgt gctcgcacct ggaatgcaga acgtttttcc tgacgcaggg agctctgctg 420
aatgacaagc attcgaatgg aacgggtcaa gacagatcgc ctacacagaac gctgatgtcg 480
tgtcccgtgg gagaagctcc ttcgccctac aactcgagat tcgaatcggg cgcttggtcg 540
gcttcggctt gtcgatgtgg aacgtcgtgg ctgacgatcg gaatctcggg acccgacaat 600
ggagctgtgg ctgtcctgaa atacaatgga atcatcacgg acacgatcaa gtcgtggaga 660
aacaacatca tgagaacgca agaatcgga tgtgcttggt tcaatggatc gtgcttcacg 720
gtcatgacgg atggaccctc gaatggacag gcttcgtaca aaatctttag aatcgaaaaa 780
ggaaaagtcg tcaaactcggc tgaactgaat gctcctaatt atcactatga agaagctcgc 840
tgttatcctg atgctggaga aatcacgtgt gtgtgcagag ataactggca tggatcgaat 900
cgaccctggg tctcgtttaa tcaaatctg gaatatcgaa tcggatatat ctgctcggga 960
gtctttggag acaatccccg cccaatgat ggaacgggat cgtgtggacc cgtgtcgcct 1020
aaaggagctt atggaatcaa aggattttcg ttcaaatcag gaaatggagt ctggatcgga 1080
agaacgaaat cgacgaattc gagatcggga ttcgaaatga tctgggatcc caatggatgg 1140
acgggaacgg actcgaattt ctcgggtcaag caagatatcg tcgctatcac ggattggctg 1200
ggatattcgg gatcgttcgt ccagcatccc gaactgacgg gactggattg catcagacct 1260
tgtttttggg tcgaactgat cagaggacga ccaaagaat cgacgatctg gacgtcggga 1320
tcgtcgatct cgttctgtgg agtcaattcg gacacgggtg gatggctcgtg gcccgcgga 1380
gctgaactgc cctttacgat cgacaagtag 1410

<210> 59
<211> 1410
<212> ADN
<213> Virus

<400> 59

atgaatacaa atcaaaaaat aataaccatt ggaacagcct gtctgatagt cggaataatt 60
agtctattat tgcagatagg agatatagtc tcgttatgga taagccattc aattcagact 120
ggagagaaaa accactctca gatatgcagt caaagtgtca ttacatatga aaacaacaca 180
tgggtgaacc aaacttatgt aaacattggc aataccaata ttgctgatgg acaggagta 240
aattcaataa tactagcggg caattcctct ctttgcccag taagtggatg ggccatatac 300
agcaaagaca atagcataag gatcgggttc aaaggagaca tttttgtcat aagagaacta 360
tttatctcat gctctcattt ggagtgcaga actttttatc tgaccaagg tgctttgctg 420

```

aatgacaagc attctaattgg aaccgtcaaa gacaggagtc cttatagaac cttaatgagc      480
tgcccgattg gtgaagctcc ttctccgtac aattcaagggt tcgaatcagt tgcttgggtca      540
gcaagtgcag gccatgacgg aatgggatgg ctgacaatcg gaatttccgg cccagataat      600
ggagcagtgg ctgttttgaa atacaatggg ataataacag atacaataaa aagtgggagg      660
aacaaaatac taagaacaca agaatcagaa tgtgtctgta taaacgggtc gtgtttcact      720
ataatgactg atggcccaag caatgggcag gcctcataca aaatattcaa aatgaagaaa      780
gggaaaatta ttaaatcagt ggagatgaat gcacctaat accactatga ggaatgctcc      840
tgtttaccctg atacaggcaa agtgggtgtgc gtgtgcagag acaattggca tgcttcgaat      900
agaccgtggg tctcttttga tcagaacctt aattatcaga tagggtacat atgtagtggg      960
gttttcgggtg ataaccgcg ttctaattgat gggagaggcg attgtggggc agtactttct     1020
aatggagcta atggagtgaaggaggattctca tttaggtatg gcaatggcgt ttggatagga     1080
agaactaaaa gcatcagctc tagaagtga tttgagatga tttgggatcc gaatggatgg     1140
acggaaaccg atagtagttt ctogataaag caggatgtta tagcattaac tgattgggtca     1200
ggatacagtg ggaactttgt ccaacatccc gaattaacag gaatgaactg cataaagcct     1260
tgtttctggg tagagttaat cagaggacag cccaaggaga gaacaatctg gactagtggg     1320
agcagcattt ctttctgtgg tgtagacagt gaaaccgcaa gctgggtcatg gccagacgga     1380
gctgatctgc cattcactat tgacaagtag                                         1410

```

<210> 60
 <211> 469
 <212> PRT
 <213> Virus

<400> 60

```

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

Val Gly Ile Ile Ser Leu Leu Leu Gln Ile Gly Asp Ile Val Ser Leu
 20              25              30

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

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

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

```

ES 2 662 030 T3

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

Val Phe Gly Asp Asn Pro Arg Ser Asn Asp Gly Arg Gly Asp Cys Gly
325 330 335

Pro Val Leu Ser Asn Gly Ala Asn Gly Val Lys Gly Phe Ser Phe Arg
340 345 350

Tyr Gly Asn Gly Val Trp Ile Gly Arg Thr Lys Ser Ile Ser Ser Arg
355 360 365

Ser Gly Phe Glu Met Ile Trp Asp Pro Asn Gly Trp Thr Glu Thr Asp
370 375 380

Ser Ser Phe Ser Ile Lys Gln Asp Val Ile Ala Leu Thr Asp Trp Ser
385 390 395 400

Gly Tyr Ser Gly Asn Phe Val Gln His Pro Glu Leu Thr Gly Met Asn
405 410 415

Cys Ile Lys Pro Cys Phe Trp Val Glu Leu Ile Arg Gly Gln Pro Lys
420 425 430

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

Asp Ser Glu Thr Ala Ser Trp Ser Trp Pro Asp Gly Ala Asp Leu Pro
450 455 460

Phe Thr Ile Asp Lys
465

<210> 61
<211> 1410
<212> ADN
<213> Virus

<400> 61

atgaatacga atcaaaaaat catcacgatac ggaacggcctt gtctgatcgt cggaatcatc	60
tcgctgctgc tgcagatcgg agatatcgtc tcgctgtgga tctcgcatc gatccagacg	120
ggagaaaaaa accactcgca gatctgctcg caatcgggtca tcacgtatga aaacaacacg	180
tgggtgaacc aaacgtatgt caacatcgga aatacgaata tcgctgatgg acagggagtc	240
aattcgatca tcttggtggtg aaattcgtcg ctgtgccccg tctcgggatg ggctatctac	300
tcgaaagaca attcgatcag aatcggatcg aaaggagaca tcttcgtcat cagagaactg	360
ttcatctcgt gctcgcatct ggaatgcaga acgttctatc tgacgcaagg agctctgctg	420

aatgacaagc attcgaatgg aacgggtcaaa gacagatcgc cttatagaac gctgatgtcg 480
 tgccccatcg gagaagctcc ttccgacctac aattcgagat ttgaatcggg cgcttggtcg 540
 gcttcggcctt gccatgacgg aatgggatgg ctgacgatcg gaatctcggg acccgataat 600
 ggagctgtgg ctgtcctgaa atacaatgga atcatcacgg atacgatcaa atcgtggaga 660
 aacaaaatcc tgagaacgca agaatcggaa tgtgtctgta tcaacggatc gtgttttacg 720
 atcatgacgg atggaccctc gaatggacag gcttcgtaca aaatctttaa aatgaagaaa 780
 ggaaaaatca tcaaatcggg ggaaatgaat gctcctaatt accactatga agaatgctcg 840
 tgttaccctg atacgggaaa agtgggtgtg gtgtgcagag acaattggca tgcttcgaat 900
 agaccctggg tctcgtttga tcagaacctg aattatcaga tcggatacat ctgttcggga 960
 gtcttttgag ataacccccg ttcgaaatgat ggaagaggag attgtggacc cgtcctgtcg 1020
 aatggagcta atggagtgaaggatcttcg ttccagatatg gaaatggagt ctggatcgga 1080
 agaacgaaat cgatctcgtc gagatcggga ttcgaaatga tctgggatcc caatggatgg 1140
 acggaaacgg attcgtcgtt ttccgatcaag caggatgtca tcgctctgac ggattggctg 1200
 ggatactcgg gaaacttcgt ccaacatccc gaactgacgg gaatgaactg catcaagcct 1260
 tgtttttgagg tcgaactgat cagaggacag cccaaggaaa gaacgatctg gacgtcggga 1320
 tcgtcgatct cgttttgtgg agtcgactcg gaaacggctt cgtggctcgtg gcccgacgga 1380
 gctgatctgc cctttacgat cgacaagtag 1410

<210> 62
 <211> 576
 <212> ADN
 <213> Virus

<400> 62

taccaagtgc gcaattcctc ggggctttac catgtcacca atgattgccc taactcgagt 60
 attgtgtacg aggcggccga tgccatcctg cacactccgg ggtgtgtccc ttgcgttcgc 120
 gagggtaacg cctcgagggtg ttgggtggcg gtgaccccca cggtgccac caggacggc 180
 aaactcccca caacgcagct tcgacgtcat atcgatctgc ttgtcgggag cgccaccctc 240
 tgctcggccc tctacgtggg ggacctgtgc gggctctgtc ttcttgttgg tcaactgttt 300
 accttctctc ccaggcgcca ctggacgacg caagactgca attgttctat ctatcccggc 360
 catataacgg gtcacgtcat ggcattggat atgatgatga actgggtccc tacggcagcg 420
 ttggtggtag ctcagctgct ccggatccca caagccatca tggacatgat cgctggtgct 480
 cactggggag tcctggcggg catagcgtat ttctccatgg tggggaactg ggcgaaggtc 540
 ctggtagtgc tgctgctatt tgccggcgtc gacgcg 576

<210> 63
 <211> 192
 <212> PRT
 <213> Virus

ES 2 662 030 T3

<400> 63

```

Tyr Gln Val Arg Asn Ser Ser Gly Leu Tyr His Val Thr Asn Asp Cys
1          5          10          15

Pro Asn Ser Ser Ile Val Tyr Glu Ala Ala Asp Ala Ile Leu His Thr
20          25          30

Pro Gly Cys Val Pro Cys Val Arg Glu Gly Asn Ala Ser Arg Cys Trp
35          40          45

Val Ala Val Thr Pro Thr Val Ala Thr Arg Asp Gly Lys Leu Pro Thr
50          55          60

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

Cys Ser Ala Leu Tyr Val Gly Asp Leu Cys Gly Ser Val Phe Leu Val
85          90          95

Gly Gln Leu Phe Thr Phe Ser Pro Arg Arg His Trp Thr Thr Gln Asp
100         105         110

Cys Asn Cys Ser Ile Tyr Pro Gly His Ile Thr Gly His Arg Met Ala
115         120         125

Trp Asp Met Met Met Asn Trp Ser Pro Thr Ala Ala Leu Val Val Ala
130         135         140

Gln Leu Leu Arg Ile Pro Gln Ala Ile Met Asp Met Ile Ala Gly Ala
145         150         155         160

His Trp Gly Val Leu Ala Gly Ile Ala Tyr Phe Ser Met Val Gly Asn
165         170         175

Trp Ala Lys Val Leu Val Val Leu Leu Leu Phe Ala Gly Val Asp Ala
180         185         190

```

5 <210> 64
 <211> 576
 <212> ADN
 <213> Virus

10 <400> 64

ES 2 662 030 T3

taccaagtgc gcaattcgtc gggactgtac catgtcacga atgattgccc taactcgtcg	60
atcgtgtacg aagctgctga tgctatcctg cacacgcccc gatgtgtccc ttgcgtccgc	120
gaaggaaacg cttcgagatg ttgggtggct gtgacgcccc cgggtggctac gagagacgga	180
aaactgcccc cgacgcagct gcgacgtcat atcgatctgc tggtcggatc ggctacgctg	240
tgctcggctc tgtacgtggg agacctgtgc ggatcggctc tcctggtcgg acaactgttc	300
acgtttttcgc ccagacgcca ctggacgacg caagactgca attgttcgat ctatcccgga	360
catatcacgg gacatcgcat ggcttgggat atgatgatga actggtcgcc tacggctgct	420
ctgggtggctg ctcagctgct gcgaatcccc caagctatca tggacatgat cgctggagct	480
cactggggag tcctggctgg aatcgcttat ttttcgatgg tgggaaactg ggctaaggtc	540
ctggctcgtc tgctgctgtt cgctggagtc gacgct	576

<210> 65
 <211> 1089
 <212> ADN
 <213> Virus

5

<400> 65

gaaacccacg tcaccggggg aagtgccggc cgcaccacgg ctgggcttgt tggctctcctt	60
acaccaggcg ccaagcagaa catccaactg atcaacacca acggcagttg gcacatcaat	120
agcacggcct tgaactgcaa tgaaagcctt aacaccggct ggtagcagg gctcttctat	180
cagcacaaat tcaactcttc aggctgtcct gagagggttg ccagctgccg acgccttacc	240
gattttgccc agggctgggg tcctatcagt tatgccaacg gaagcggcct cgacgaacgc	300
ccctactgct ggcactaccc tccaagacct tgtggcattg tgcccgcaa gagcgtgtgt	360
ggcccgggat attgcttcac tcccagcccc gtgggtggtg gaacgaccga caggctgggc	420
gcgcctacct acagctgggg tgcaaatgat acggatgtct tcgtccttaa caacaccagg	480
ccaccgctgg gcaattgggt cggttggtacc tggatgaact caactggatt caccaaagtg	540
tgcggagcgc ccccttggtt catcggaggg gtgggcaaca acaccttgct ctgcccact	600
gattgtttcc gcaagcatcc ggaagccaca tactctcgtt gcggctccgg tccttggtt	660
acaccaggt gcatggtcga ctaccctgat aggctttggc actatccttg taccatcaat	720
tacaccatat tcaaagtcag gatgtacgtg ggaggggtcg agcacaggct ggaagcggcc	780
tgcaactgga cgcggggcga acgctgtgat ctggaagaca gggacaggtc cgagctcagc	840
ccattgctgc tgtccaccac acagtggcag gtccctccgt gttctttcac gacctgcca	900
gccttggtcca ccggcctcat ccacctccac cagaacattg tggacgtgca gtactgttac	960
ggggtagggc caagcatcgc gtccctgggc attaatggg agtacgtcgt tctcctgttc	1020
ctcctgcttg cagacgcgcg cgtctgctcc tgcttggtga tgatgttact catatcccaa	1080
gcggaggcgc	1089

10

<210> 66
 <211> 363
 <212> PRT
 <213> Virus

15

ES 2 662 030 T3

<400> 66

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

ES 2 662 030 T3

Asn Asn Thr Leu Leu Cys Pro Thr Asp Cys Phe Arg Lys His Pro Glu
 195 200 205
 Ala Thr Tyr Ser Arg Cys Gly Ser Gly Pro Trp Ile Thr Pro Arg Cys
 210 215 220
 Met Val Asp Tyr Pro Tyr Arg Leu Trp His Tyr Pro Cys Thr Ile Asn
 225 230 235 240
 Tyr Thr Ile Phe Lys Val Arg Met Tyr Val Gly Gly Val Glu His Arg
 245 250 255
 Leu Glu Ala Ala Cys Asn Trp Thr Arg Gly Glu Arg Cys Asp Leu Glu
 260 265 270
 Asp Arg Asp Arg Ser Glu Leu Ser Pro Leu Leu Leu Ser Thr Thr Gln
 275 280 285
 Trp Gln Val Leu Pro Cys Ser Phe Thr Thr Leu Pro Ala Leu Ser Thr
 290 295 300
 Gly Leu Ile His Leu His Gln Asn Ile Val Asp Val Gln Tyr Leu Tyr
 305 310 315 320
 Gly Val Gly Ser Ser Ile Ala Ser Trp Ala Ile Lys Trp Glu Tyr Val
 325 330 335
 Val Leu Leu Phe Leu Leu Leu Ala Asp Ala Arg Val Cys Ser Cys Leu
 340 345 350
 Trp Met Met Leu Leu Ile Ser Gln Ala Glu Ala
 355 360

<210> 67
 <211> 1089
 <212> ADN
 <213> Virus

<400> 67

gaaacgcacg tcacgggagg atcggtgga cgcacgacg ctggactggt cggactgctg	60
acgcccggag ctaagcagaa catccaactg atcaacacga acggatcgtg gcacatcaat	120
tcgacggctc tgaactgcaa tgaatcgctg aacacgggat ggctggctgg actgttttat	180
cagcaciaaat ttaactcgtc gggatgtcct gaaagactgg cttcgtgccg acgcctgacg	240
gatttcgctc agggatgggg acctatctcg tatgctaacy gatcgggact ggacgaacgc	300
ccctactgct ggcactaccc tcccagacct tgtggaatcg tgcccgctaa gtcggtgtgt	360

ggacccgtct attgctttac gccctcgccc gtggtggtgg gaacgacgga cagatcgggg 420
gctcctacgt actcgtgggg agctaagat acggatgtct ttgtcctgaa caacacgaga 480
ccccccctgg gaaattggtt tggatgtacg tggatgaact cgacgggatt tacgaaagtg 540
tgcggagctc ccccttgtgt catcggagga gtgggaaaca acacgctgct gtgccccacg 600
gattgttttc gcaagcatcc cgaagctacg tactcgcgat gcggatcggg accctggatc 660
acgcccagat gcatggtcga ctacccttat agactgtggc actatccttg tacgatcaat 720
tacacgatct ttaaagtcag aatgtacgtg ggaggagtcg aacacagact ggaagctgct 780
tgcaactgga cgcgaggaga acgctgtgat ctggaagaca gagacagatc ggaactgtcg 840
cccctgctgc tgtcgacgac gcagtggcag gtctcgccct gttcgtttac gacgctgccc 900
gctctgtcga cgggactgat ccacctgcac cagaacatcg tggacgtgca gtacctgtac 960
ggagtcggat cgtcgatcgc ttcgtgggct atcaagtggg aatacgtcgt cctgctgttt 1020
ctgctgctgg ctgacgctcg cgtctgctcg tgcctgtgga tgatgctgct gatctcgaa 1080
getgaagct 1089

<210> 68
<211> 2724
<212> ADN
<213> Virus

<400> 68

atggaggcag ccttgcttgt gtgtcagtac accatccaga gcctgatcca tctcacgggt 60
gaagatcctg gttttttcaa tgttgagatt ccggaattcc cattttaccc cacatgcaat 120
gtttgcacgg cagatgtcaa tgtaactatc aatttcgatg tcggggggcaa aaagcatcaa 180
cttgatcttg actttggcca gctgacaccc catacgaagg ctgtctacca acctcgaggt 240
gcatttggtg gctcagaaaa tgccaccaat ctctttctac tggagctcct tgggtgcagga 300
gaattggctc taactatgcg gtctaagaag cttccaatta acgtcaccac cggagaggag 360
caacaagtaa gcctggaatc tgtagatgtc tactttcaag atgtgttttg aaccatgtgg 420
tgccaccatg cagaaatgca aaaccccggt tacctgatac cagaaacagt gccatacata 480
aagtgggata actgtaattc taccaatata acggcagtag tgagggcaca ggggctggat 540
gtcacgctac ccttaagttt gccaacgtca gctcaagact cgaatttcag cgtaaaaaca 600
gaaatgctcg gtaatgagat agatattgag tgtattatgg aggatggcga aatttcacaa 660
gttctgcccg gagacaacaa atttaacatc acctgcagtg gatacgagag ccatgttccc 720
agcggcggaa ttctcacatc aacgagtcct gtggccaccc caatacctgg tacagggtat 780
gcatacagcc tgcgtctgac accacgtcca gtgtcacgat ttcttggtgaa taacagtatc 840

ES 2 662 030 T3

```

ctgtacgtgt tttactctgg gaatggaccg aaggcgagcg ggggagatta ctgcattcag      900
tccaacattg tgttctctga tgagattcca gcttcacagg acatgccgac aaacaccaca      960
gacatcacat atgtgggtga caatgctacc tattcagtgc caatggtcac ttctgaggac     1020
gcaaactcgc caaatgttac agtgactgcc ttttgggcct ggccaaacaa cactgaaact     1080
gactttaagt gcaaatggac tctcacctcg gggacacctt cgggttgtga aaatatctct     1140
ggtgcatttg cgagcaatcg gacatttgac attactgtct cgggtcttgg caccgcccc     1200
aagacactca ttatcacacg aacggctacc aatgccacca caacaacca caaggttata     1260
tttccaagg cacccgagag caccaccacc tcccctacct tgaatacaac tggatttgct     1320
gatcccaata caacgacagg tctaccagc tctactcacg tgcctacca cctcaccgca     1380
cctgcaagca caggccccac tgtatccacc gcggtatgtca ccagcccaac accagccggc     1440
acaacgtcag gcgcatcacc ggtgacacca agtccatctc catgggacaa cggcacagaa     1500
agtaaggccc ccgacatgac cagctccacc tcaccagtga ctacccaac cccaaatgcc     1560
accagcccca cccagcagt gactaccca accccaaatg ccaccagccc caccagca     1620
gtgactacce caacccaaa tgccaccagc cccaccttgg gaaaaacaag tcctacctca     1680
gcagtgacta ccccaacccc aaatgccacc agccccacct tgggaaaaac aagccccacc     1740
tcagcagtga ctacccaac cccaaatgcc accagcccca ccttgggaaa aacaagcccc     1800
acctcagcag tgactacccc aaccccaaat gccaccggcc ctactgtggg agaaacaagt     1860
ccacaggcaa atgccacca ccacacctta ggaggaacaa gtccacccc agtagttacc     1920
agccaaccaa aaaatgcaac cagtgtgttt accacaggcc aacataacat aacttcaagt     1980
tcaacctctt ccatgtcact gagaccagt tcaaaccag agacactcag cccctccacc     2040
agtgacaatt caacgtcaca tatgccttta ctaacctcg ctcaccaac aggtggtgaa     2100
aatataacac aggtgacacc agcctctatc agcacacatc atgtgtccac cagttcgcca     2160
gcaccccgcc caggcaccac cagccaagcg tcaggccctg gaaacagttc cacatccaca     2220
aaacgggggg aggttaatgt caccaaaggc acgccccccc aaaatgcaac gtcgccccag     2280
gccccagtg gccaaaagac ggcggttccc acggtcacct caacaggtgg aaaggccaat     2340
tctaccaccg gtggaaagca caccacagga catggagccc ggacaagtac agagcccacc     2400
acagattacg gcggtgattc aactacgcca agaccgagat acaatgcgac cacctatcta     2460
cctcccagca cttctagcaa actgcccgcg cgctggactt ttacgagccc accggttacc     2520
acagcccaag ccaccgtgcc agtcccgcga acgtcccagc ccagattctc aaacctctcc     2580
atgctagtac tgcagtgggc ctctctggct gtgctgaccc ttctgctgct gctggtcatg     2640

gcggactgcg cctttaggcg taacttgtct acatcccata cctacaccac cccacatat      2700
gatgacgccg agacctatgt ataa                                         2724

```

<210> 69
 <211> 907
 <212> PRT
 <213> Virus

<400> 69

ES 2 662 030 T3

[illegible]

ES 2 662 030 T3

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

ES 2 662 030 T3

Pro Ser Ser Thr His Val Pro Thr Asn Leu Thr Ala Pro Ala Ser Thr
 450 455 460

Gly Pro Thr Val Ser Thr Ala Asp Val Thr Ser Pro Thr Pro Ala Gly
 465 470 475 480

Thr Thr Ser Gly Ala Ser Pro Val Thr Pro Ser Pro Ser Pro Trp Asp
 485 490 495

Asn Gly Thr Glu Ser Lys Ala Pro Asp Met Thr Ser Ser Thr Ser Pro
 500 505 510

Val Thr Thr Pro Thr Pro Asn Ala Thr Ser Pro Thr Pro Ala Val Thr
 515 520 525

Thr Pro Thr Pro Asn Ala Thr Ser Pro Thr Pro Ala Val Thr Thr Pro
 530 535 540

Thr Pro Asn Ala Thr Ser Pro Thr Leu Gly Lys Thr Ser Pro Thr Ser
 545 550 555 560

Ala Val Thr Thr Pro Thr Pro Asn Ala Thr Ser Pro Thr Leu Gly Lys
 565 570 575

Thr Ser Pro Thr Ser Ala Val Thr Thr Pro Thr Pro Asn Ala Thr Ser
 580 585 590

Pro Thr Leu Gly Lys Thr Ser Pro Thr Ser Ala Val Thr Thr Pro Thr
 595 600 605

Pro Asn Ala Thr Gly Pro Thr Val Gly Glu Thr Ser Pro Gln Ala Asn
 610 615 620

Ala Thr Asn His Thr Leu Gly Gly Thr Ser Pro Thr Pro Val Val Thr
 625 630 635 640

Ser Gln Pro Lys Asn Ala Thr Ser Ala Val Thr Thr Gly Gln His Asn
 645 650 655

Ile Thr Ser Ser Ser Thr Ser Ser Met Ser Leu Arg Pro Ser Ser Asn
 660 665 670

Pro Glu Thr Leu Ser Pro Ser Thr Ser Asp Asn Ser Thr Ser His Met
 675 680 685

ES 2 662 030 T3

Pro Leu Leu Thr Ser Ala His Pro Thr Gly Gly Glu Asn Ile Thr Gln
 690 695 700
 Val Thr Pro Ala Ser Ile Ser Thr His His Val Ser Thr Ser Ser Pro
 705 710 715 720
 Ala Pro Arg Pro Gly Thr Thr Ser Gln Ala Ser Gly Pro Gly Asn Ser
 725 730 735
 Ser Thr Ser Thr Lys Pro Gly Glu Val Asn Val Thr Lys Gly Thr Pro
 740 745 750
 Pro Gln Asn Ala Thr Ser Pro Gln Ala Pro Ser Gly Gln Lys Thr Ala
 755 760 765
 Val Pro Thr Val Thr Ser Thr Gly Gly Lys Ala Asn Ser Thr Thr Gly
 770 775 780
 Gly Lys His Thr Thr Gly His Gly Ala Arg Thr Ser Thr Glu Pro Thr
 785 790 795 800
 Thr Asp Tyr Gly Gly Asp Ser Thr Thr Pro Arg Pro Arg Tyr Asn Ala
 805 810 815
 Thr Thr Tyr Leu Pro Pro Ser Thr Ser Ser Lys Leu Arg Pro Arg Trp
 820 825 830
 Thr Phe Thr Ser Pro Pro Val Thr Thr Ala Gln Ala Thr Val Pro Val
 835 840 845
 Pro Pro Thr Ser Gln Pro Arg Phe Ser Asn Leu Ser Met Leu Val Leu
 850 855 860
 Gln Trp Ala Ser Leu Ala Val Leu Thr Leu Leu Leu Leu Val Met
 865 870 875 880
 Ala Asp Cys Ala Phe Arg Arg Asn Leu Ser Thr Ser His Thr Tyr Thr
 885 890 895
 Thr Pro Pro Tyr Asp Asp Ala Glu Thr Tyr Val
 900 905

<210> 70
 <211> 2724
 <212> ADN
 <213> Virus

<400> 70

atggaagctg ctctgctggt gtgtcagtac acgatccagt cgctgatcca tctgacggga	60
gaagatccctg gattctttaa tgtcgaaatc cccgaatttc ccttctaccc cacgtgcaat	120
gtctgcacgg ctgatgtcaa tgtcacgac aattttgatg tcggaggaaa aaagcatcaa	180
ctggatctgg acttcggaca gctgacgccc catacgaagg ctgtctacca acctcgagga	240
gctttcggag gatcgaaaa tgctacgaat ctgttcctgc tggaactgct gggagctgga	300
gaactggctc tgacgatgag atcgaagaag ctgccatca acgtcacgac gggagaagaa	360
caacaagtct cgctggaatc ggtcgatgtc tacttccaag atgtgttcgg aacgatgtgg	420
tgccaccatg ctgaaatgca aaaccccggt tactctgatcc ccgaaacggt gccctacatc	480
aagtgggata actgtaattc gacgaatatc acggctgtcg tgagagctca gggactggat	540
gtcacgctgc cctgtcgtc gccacgctg gctcaagact cgaatttttc ggtcaaaacg	600
gaaatgctgg gaaatgaaat cgatatcgaa tgtatcatgg aagatggaga aatctcgcaa	660
gtcctgcccc gagacaacaa attcaacatc acgtgctcgg gatacgaatc gcatgtcccc	720
tcgggaggaa tcctgacgtc gacgtcgccc gtggctacgc ccatccctgg aacgggatat	780
gcttactcgc tgcgtctgac gcccgctccc gtgtcgcgat tcctgggaaa taactcgatc	840
ctgtacgtgt tctactcggg aaatggaccc aaggttcgg gaggagatta ctgcatccag	900
tcgaacatcg tgttttcgga tgaaatcccc gcttcgcagg acatgccac gaacacgacg	960
gacatcacgt atgtgggaga caatgctacg tattcgggtgc ccatggtcac gtcggaagac	1020
gctaactcgc ccaatgtcac ggtgacggct ttctgggctt ggcccaacaa cacggaaacg	1080
gacttcaagt gcaaatggac gctgacgtcg ggaacgcctt cgggatgtga aaatatctcg	1140
ggagctttcg cttcgaatcg aacgttcgac atcacggctc cgggactggg aacggctccc	1200
aagacgctga tcatcacgag aacggctacg aatgctacga cgacgacgca caaggctcatc	1260
ttttcgaagg ctcccgaatc gacgacgacg tcgcctacgc tgaatacgac gggattcgtc	1320
gatcccaata cgacgacggg actgccctcg tcgacgcacg tgcctacgaa cctgacggct	1380
cctgcttcga cgggacccac ggtctcgacg gctgatgtca cgtcgccac gccgctgga	1440
acgacgtcgg gagcttcgcc cgtgacgccc tcgcctcgc cctgggacaa cggaacggaa	1500
tcgaaggctc ccgacatgac gtgctcgacg tcgcccgtga cgacgcccac gcccaatgct	1560
acgtcgccca cggcgtgtg gacgacgccc acgcccattg ctacgtcgcc cagcccgtc	1620
gtgacgacgc ccacgcccac tgctacgtcg cccacgctgg gaaaaacgtc gcctacgtcg	1680
gctgtgacga cggccacgcc caatgctacg tcgcccacgc tgggaaaaac gtcgcccacg	1740
tcggctgtga cgacgcccac gcccaatgct acgtcgccca cgctgggaaa aacgtcgccc	1800

acgtcggctg tgacgacgcc cagcccaat gctacgggac ctacggtggg agaaacgtcg 1860
 cccagggcta atgctacgaa ccacacgctg ggaggaacgt cgcacacgcc cgtcgtcacg 1920
 tcgcaaccca aaaatgctac gtcggctgtc acgacgggac aacataacat cagctcgtcg 1980
 tcgacgtcgt cgatgtcgt gagaccctcg tcgaaccccg aaacgctgtc gccctcgacg 2040
 tcggacaatt cgacgtcgca tatgcctctg ctgacgtcgg ctcacccacac gggaggagaa 2100
 aatatcacgc aggtgacgcc cgcttcgatc tcgacgcac atgtgtcgac gtcgtcgccc 2160
 gctccccgcc ccggaacgac gtcgcaagct tcgggacctg gaaactcgtc gacgtcgacg 2220
 aaaccggag aagtcaatgt cagaaagga acgcccccc aaaatgctac gtcgccccag 2280
 gctccctcgg gacaaaagac ggctgtcccc acggtcacgt cgacgggagg aaaggcta 2340
 tcgacgacgg gaggaaagca cagcagggga catggagctc gaacgtcgac ggaacccacg 2400
 acggattacg gaggagattc gacgacgccc agaccagat acaatgctac gacgtatctg 2460
 cctccctcga cgtcgtcgaa actgcgaccc cgtgggacgt tcacgtcgcc ccccgtcacg 2520
 acggctcaag ctacggtgcc cgtcccccc acgtcgcagc ccagattttc gaacctgtcg 2580
 atgctgggtc tgcagtgggc ttctgtggct gtgtgacgc tgctgctgct gctggtcagt 2640
 gctgactcgc ctttcagacg taacctgtcg acgtcgcata cgtacacgac gccccctat 2700
 gatgacgctg aaacgtatgt ctaa 2724

<210> 71
 <211> 2661
 <212> ADN
 <213> Virus

<400> 71

atggaggcag ccttgcttgt gtgtcagtac accatccaga gccttatcca actcacgcgt 60
 gatgatcctg gttttttcaa tgttgagatt ctggaattcc cattttaccc agcgtgcaat 120
 gtttgacgag cagatgtcaa tgcaactatc aatttcgatg tcgggggcaa aaagcataaa 180
 cttaatcttg actttggcct gctgacaccc catacaaagg ctgtctacca acctcgaggt 240
 gcatttggtg gtcagaaaaa tgccaccaat ctctttctac tggagctcct tgggtgcagga 300
 gaattggctc taactatgag gtctaagaag cttccaatta acatcacccac cggagaggag 360
 caacaagtaa gcctggaatc tgtagatgtc tactttcaag atgtgttttg caccatgtgg 420
 tgccaccatg cagaaatgca aaaccagta tacctaatac cagaaacagt gccatacata 480
 aagtgggata actgtaattc taccaatata acggcagtag taagggcaca ggggctggat 540
 gtcacgctac ccttaagttt gccaacatca gctcaagact cgaatttcag cgtaaaaaa 600
 gaaatgctcg gtaatgagat agatattgag tgtattatgg aggatggcga aatttcacaa 660

```

gttctgcccc gagacaacaa atttaacatc acctgcagtg gatacgagag ccatgttccc 720
agcggcgga ttctcacatc aacgagtcct gtggccaccc caatacctgg tacagggtat 780
gcatacagcc tgcgtctgac accacgtcca gtgtcacgat ttcttgga taacagtata 840
ctgtacgtgt ttactctgg gaattggacc aaggcgagcg ggggagatta ctgcattcag 900
tccaacattg tgttctctga tgagattcca gcttcacagg acatgccgac aaacaccaca 960
gacatcacat atgtgggtga caatgctacc tattcagtgc caatgggtcac ttctgaggac 1020
gcaaaactgc caaatgttac agtgactgcc ttttgggcct ggccaaacaa cactgaaact 1080
gactttaagt gcaaatggac tctcacctcg gggacacctt cgggttgtga aaatatttct 1140
ggtgcatttg cgagcaatcg gacatttgac attactgtct cgggtcttgg caccgcccc 1200
aagacactca ttatcacacg aacggctacc aatgccacca caacaacca caaggtata 1260
ttctccaagg caccgagag caccaccacc tccctacct tgaatacaac tggatttgc 1320
gtcccaata caacgacagg tctaccagc tctactcag tgcctacca cctaccgca 1380
cctgcaagca caggccccac tgtatccacc gcggatgtca ccagcccaac accagccggc 1440
acaacgtcag gcgcatcacc ggtgacacca agtccatctc caccgggacaa cggcacagaa 1500
agtaaggccc ccgacatgac cagccccacc tcagcagtga ctacccaac cccaaatgcc 1560
accagcccca cccagcagt gactaccca accccaaatg ccaccagccc cacttggga 1620
aaaacaagtc ccacctcagc agtgactacc ccaaccccaa atgccaccag cccacccca 1680
gcagtgacta ccccaacccc aaatgccacc atccccacct tgggaaaaac aagtcacc 1740
tcagcagtga ctacccaac cccaaatgcc accagcccta ccgtgggaga aacaagtcca 1800
caggcaata ccaccaacca cacattagga ggaacaagtt ccacccagc agttaccagc 1860
ccacaaaaa atgcaaccag tgctgttacc acaggccaa ataacataac ttcaagttca 1920
acctctcca tgtcactgag acccagttca atctcagaga cactcagccc ctccaccag 1980
gacaattcaa cgtcacatat gcctttacta acctccgctc acccaacagg tggtgaaaat 2040
ataacacagg tgacaccagc ctctaccagc acacatcatg tgtccaccag ttccgacagc 2100
ccccgcccag gcaccaccag ccaagcgtca ggccctggaa acagttccac atccaaaaa 2160
ccgggggagg ttaatgtcac caaaggcagc cccccaaaa atgcaacgtc gccccaggcc 2220
cccagtggcc aaaagacggc ggttcccagc gtcacctcaa cagggtgaaa ggccaattct 2280
accaccggtg gaaagcacac cacaggacat ggagcccgga caagtacaga gccaccaca 2340
gattacggcg gtgattcaac tacgccaaga acgagatata atgcgaccac ctatctacct 2400
ccagcactt ctagaaaact ggggccccgc tggactttta cgagccacc gggtaccaca 2460
gccaagcca ccgtgcctgt cccgccaacg tccagccca gattctcaa cctctccatg 2520
ctagtactgc agtgggcctc tctggctgtg ctgaccttc tgctgtgtgt ggtcatggcg 2580
gactgcgctt tcaggcgtaa cttgtcgaca tccataacct acaccacccc accatatgat 2640
gacgccgaga cctatgtata a 2661

```

5

<210> 72
 <211> 886
 <212> PRT
 <213> Virus

ES 2 662 030 T3

<400> 72

```

Met Glu Ala Ala Leu Leu Val Cys Gln Tyr Thr Ile Gln Ser Leu Ile
1          5          10          15

Gln Leu Thr Arg Asp Asp Pro Gly Phe Phe Asn Val Glu Ile Leu Glu
          20          25          30

Phe Pro Phe Tyr Pro Ala Cys Asn Val Cys Thr Ala Asp Val Asn Ala
          35          40          45

Thr Ile Asn Phe Asp Val Gly Gly Lys Lys His Lys Leu Asn Leu Asp
          50          55          60

Phe Gly Leu Leu Thr Pro His Thr Lys Ala Val Tyr Gln Pro Arg Gly
65          70          75          80

Ala Phe Gly Gly Ser Glu Asn Ala Thr Asn Leu Phe Leu Leu Glu Leu
          85          90          95

Leu Gly Ala Gly Glu Leu Ala Leu Thr Met Arg Ser Lys Lys Leu Pro
          100          105          110

Ile Asn Ile Thr Thr Gly Glu Glu Gln Gln Val Ser Leu Glu Ser Val
          115          120          125

Asp Val Tyr Phe Gln Asp Val Phe Gly Thr Met Trp Cys His His Ala
          130          135          140

Glu Met Gln Asn Pro Val Tyr Leu Ile Pro Glu Thr Val Pro Tyr Ile
145          150          155          160

Lys Trp Asp Asn Cys Asn Ser Thr Asn Ile Thr Ala Val Val Arg Ala
          165          170          175

Gln Gly Leu Asp Val Thr Leu Pro Leu Ser Leu Pro Thr Ser Ala Gln

```

ES 2 662 030 T3

[illegible]

ES 2 662 030 T3

Thr	Leu	Asn	Thr	Thr	Gly	Phe	Ala	Ala	Pro	Asn	Thr	Thr	Thr	Gly	Leu			
		435					440					445						
Pro	Ser	Ser	Thr	His	Val	Pro	Thr	Asn	Leu	Thr	Ala	Pro	Ala	Ser	Thr			
		450					455				460							
Gly	Pro	Thr	Val	Ser	Thr	Ala	Asp	Val	Thr	Ser	Pro	Thr	Pro	Ala	Gly			
465					470					475					480			
Thr	Thr	Ser	Gly	Ala	Ser	Pro	Val	Thr	Pro	Ser	Pro	Ser	Pro	Arg	Asp			
			485						490					495				
Asn	Gly	Thr	Glu	Ser	Lys	Ala	Pro	Asp	Met	Thr	Ser	Pro	Thr	Ser	Ala			
			500					505					510					
Val	Thr	Thr	Pro	Thr	Pro	Asn	Ala	Thr	Ser	Pro	Thr	Pro	Ala	Val	Thr			
		515					520					525						
Thr	Pro	Thr	Pro	Asn	Ala	Thr	Ser	Pro	Thr	Leu	Gly	Lys	Thr	Ser	Pro			
	530					535					540							
Thr	Ser	Ala	Val	Thr	Thr	Pro	Thr	Pro	Asn	Ala	Thr	Ser	Pro	Thr	Pro			
545					550					555					560			
Ala	Val	Thr	Thr	Pro	Thr	Pro	Asn	Ala	Thr	Ile	Pro	Thr	Leu	Gly	Lys			
				565					570					575				
Thr	Ser	Pro	Thr	Ser	Ala	Val	Thr	Thr	Pro	Thr	Pro	Asn	Ala	Thr	Ser			
			580					585					590					
Pro	Thr	Val	Gly	Glu	Thr	Ser	Pro	Gln	Ala	Asn	Thr	Thr	Asn	His	Thr			
		595					600					605						
Leu	Gly	Gly	Thr	Ser	Ser	Thr	Pro	Val	Val	Thr	Ser	Pro	Pro	Lys	Asn			
	610					615					620							
Ala	Thr	Ser	Ala	Val	Thr	Thr	Gly	Gln	His	Asn	Ile	Thr	Ser	Ser	Ser			
625					630					635					640			
Thr	Ser	Ser	Met	Ser	Leu	Arg	Pro	Ser	Ser	Ile	Ser	Glu	Thr	Leu	Ser			
			645						650					655				
Pro	Ser	Thr	Ser	Asp	Asn	Ser	Thr	Ser	His	Met	Pro	Leu	Leu	Thr	Ser			
			660					665					670					

ES 2 662 030 T3

Ala His Pro Thr Gly Gly Glu Asn Ile Thr Gln Val Thr Pro Ala Ser
675 680 685

Thr Ser Thr His His Val Ser Thr Ser Ser Pro Ala Pro Arg Pro Gly
690 695 700

Thr Thr Ser Gln Ala Ser Gly Pro Gly Asn Ser Ser Thr Ser Thr Lys
705 710 715 720

Pro Gly Glu Val Asn Val Thr Lys Gly Thr Pro Pro Lys Asn Ala Thr
725 730 735

Ser Pro Gln Ala Pro Ser Gly Gln Lys Thr Ala Val Pro Thr Val Thr
740 745 750

Ser Thr Gly Gly Lys Ala Asn Ser Thr Thr Gly Gly Lys His Thr Thr
755 760 765

Gly His Gly Ala Arg Thr Ser Thr Glu Pro Thr Thr Asp Tyr Gly Gly
770 775 780

Asp Ser Thr Thr Pro Arg Thr Arg Tyr Asn Ala Thr Thr Tyr Leu Pro
785 790 795 800

Pro Ser Thr Ser Ser Lys Leu Arg Pro Arg Trp Thr Phe Thr Ser Pro
805 810 815

Pro Val Thr Thr Ala Gln Ala Thr Val Pro Val Pro Pro Thr Ser Gln
820 825 830

Pro Arg Phe Ser Asn Leu Ser Met Leu Val Leu Gln Trp Ala Ser Leu
835 840 845

Ala Val Leu Thr Leu Leu Leu Leu Leu Val Met Ala Asp Cys Ala Phe
850 855 860

Arg Arg Asn Leu Ser Thr Ser His Thr Tyr Thr Thr Pro Pro Tyr Asp
865 870 875 880

Asp Ala Glu Thr Tyr Val
885

<210> 73
<211> 2661
<212> ADN
<213> Virus

<400> 73

ES 2 662 030 T3

atggaagctg ctctgctggt gtgtcagtac acgatccagt cgctgatcca actgacgcgt	60
gatgatcctg gattctttaa tgtcgaaatc ctggaatttc ccttctaccc cgcttgcaat	120
gtctgcacgg ctgatgtcaa tgctacgata aattttgatg tcggaggaaa aaagcataaa	180
ctgaatctgg acttcggact gctgacgccc catacgaagg ctgtctacca acctcgagga	240
gctttcggag gatcggaaaa tgctacgaat ctgttcctgc tggaaactgct gggagctgga	300
gaactggctc tgacgatgcg atcgaagaag ctgccatca acatcacgac gggagaagaa	360
caacaagtct cgctggaatc ggtcgatgtc tacttccaag atgtgttctg aacgatgtgg	420
tgccaccatg ctgaaatgca aaaccccgtc tacctgatcc ccgaaacggt gccctacatc	480
aagtgggata actgtaattc gacgaatatc acggctgtcg tcagagctca gggactggat	540
gtcacgctgc cctgtcgct gccacgctg gctcaagact cgaatttttc ggtcaaaacg	600
gaaatgctgg gaaatgaaat cgatatcgaa tgtatcatgg aagatggaga aatctcgcaa	660
gtcctgcccg gagacaacaa attcaacatc acgtgctcgg gatacgaatc gcatgtcccc	720
tcgggaggaa tcctgacgtc gacgtcgccc gtggctacgc ccacccctgg aacgggatat	780
gcttactcgc tgcgtctgac gcccgtccc gtgtcgcgat tcctgggaaa taactcgatc	840
ctgtacgtgt tctactcggg aaatggaccc aaggcttcgg gaggagatta ctgcatccag	900
tcgaacatcg tgttttcgga tgaatcccc gcttcgcagg acatgccac gaacacgacg	960
gacatcacgt atgtgggaga caatgctacg tattcggtgc ccattggtcac gtcggaagac	1020
gctaactcgc ccaatgtcac ggtgacggct ttctgggctt ggccaacaa cacggaaacg	1080
gaacttcaagt gcaaatggac gctgacgtcg ggaacgcctt cgggatgtga aaatatctcg	1140
ggagctttcg cttcgaatcg aacgttcgac atcacggtct cgggactggg aacggctccc	1200
aagacgctga tcatcacgcg aacggctacg aatgctacga cgacgacgca caaggctcatc	1260
ttttcgaagg ctcccgaatc gacgacgacg tcgcctacgc tgaatacgac gggattcgct	1320
gtcccaata cgacgacggg actgccctcg tcgacgcacg tgcctacgaa cctgacggct	1380
cctgcttcga cgggacccac ggtctcgacg gctgatgtca cgtcgccac gcccgctgga	1440
acgacgtcgg gagcttcgcc cgtgacgccc tcgccctcgc cccgagacaa cggaacggaa	1500
tcgaaggtc cgcacatgac gtcgcccacg tcggctgtga cgacgccac gcccaatgct	1560
acgtcgcca cggccgctgt gacgacgccc acgccaatg ctacgtcgcc cacgctggga	1620
aaaacgtcgc ccacgtcggc tgtgacgacg cccacgcca atgctacgtc gccacgccc	1680
gctgtgacga cggccacgcc caatgctacg atccccacgc tgggaaaaac gtcgcccacg	1740
tcggctgtga cgacgccac gcccaatgct acgtcgcta cggtaggaga aacgtcgccc	1800

ES 2 662 030 T3

caggctaata cgacgaacca cacgctggga ggaacgtcgt cgacgcccgt cgtcacgtcg	1860
cccccaaaa atgctacgtc ggctgtcacg acgggacaac ataacatcac gtcgtcgtcg	1920
acgtcgtcga tgcgtcgtgag accctcgtcg atctcggaaa cgctgtcgcc ctcgacgtcg	1980
gacaattcga cgtcgcataat gcctctgctg acgtcggctc accccacggg aggagaaaat	2040
atcacgcagg tgacgcccgc ttcgacgtcg acgcatcatg tgcgcacgtc gtcgcccgt	2100
ccccgcccc gaacgacgtc gcaagcttcg ggacctggaa actcgtcgac gtcgacgaaa	2160
cccggagaag tcaatgtcac gaaaggaacg cccccaaaa atgctacgtc gcccaggt	2220
ccctcgggac aaaagacggc tgtccccacg gtcacgtcga cgggaggaaa ggctaattcg	2280
acgacgggag gaaagcacac gacgggacat ggagctcgaa cgtcgacgga acccacgacg	2340
gattacggag gagattcgac gacgcccaga acgagataca atgctacgac gtatctgcct	2400
ccctcgcagc cgtcgaaact gcgaccccgc tggacgttca cgtcgcccc cgtcacgacg	2460
gctcaagcta cggcgctgt cccccccacg tcgcagccca gattttcgaa cctgtcgatg	2520
ctggctcctgc agtgggcttc gctgggtgtg ctgacgtcgc tgctgctgct ggtcatggct	2580
gactgcgctt ttagacgtaa cctgtcgacg tcgcatacgt acacgacgcc cccctatgat	2640
gacgctgaaa cgtatgtcta a	2661

<210> 74
 <211> 2715
 <212> ADN
 <213> Virus

<400> 74

atgcgcgggg ggggcttgat ttgcgcgctg gtcgtggggg cgctgggtggc cgcgggtggcg	60
tcggcgggccc cggcgggccc ggcgggcccc cgcgcctcgg gcggcggtggc cgcgaccgtc	120
gcggcgaaacg ggggtcccg ctcggggccc cccccgctc cgagccccgc gaccaccaag	180
gcccgggaagc ggaaaaccaa aaagccgccc aagcggcccc aggcgacccc gcccccgac	240
gccaacgcga ccgtcgccgc cggccacgcc acgctgcgcg cgcacctgcg ggaaatcaag	300
gtcgagaacg ccgatgccca gttttacgtg tgcccgcccc cgacggggcg caccggtggcg	360
cagtttgagc agccgcgccg ctgcccgcgc cgcgggagg ggcagaacta caccgagggc	420
atcgcggtgg tcttcaagga gaacatcgcc ccgtacaaat tcaaggccac catgtactac	480
aaagacgtga ccgtgtcgca ggtgtgggtc ggccaccgct actcccagtt tatggggata	540
ttcgaggacc gcgccccgt tcccttcgag gaggtgatcg acaagattaa caccaagggg	600
gtctgccgct ccacggccaa gtacgtgcgc aacaacatgg agaccaccgc gtttcaccgg	660
gacgaccacg agaccgacat ggagctcaag ccggcggaagg tcgccacgcg caccgagccgg	720

```

gggtggcaca ccaccgacct caagtacaac ccctcgcggtg tggaggcggt ccatcggtac      780
ggcacgacgg tcaactgcat cgtcgaggag gtggacgcgc ggtcggtgta cccgtacgat      840
gagtttgtgc tggcgacggg cgactttgtg tacatgtccc cgttttacgg ctaccgggag      900
gggtcgacaca ccgagcacac cagctacgcc gccgaccgct tcaagcaggt cgacggcttc      960
tacgcgcgcg acctcaccac gaaggcccg gccacgtcgc cgacgacccg caacttgtgt      1020
acgacccccca agtttaccgt ggcctgggac tgggtgcgga agcgacccgc ggtctgcacc      1080
atgaccaagt ggcaggaggt ggacgagatg ctccgcgcgc agtacggcgc ctccctccgc      1140
ttctcctccg acgccatctc gaccaccttc accaccaacc tgaccgagta ctgcctctcg      1200
cgcgctcgacc tgggcgactg catcgcccg gatgcccgcg aggccatcga ccgcatgttt      1260
gcgcgcaagt acaacgccac gcacatcaag gtgggccagc cgcagtacta cctggccacg      1320
gggggcttcc tcatcgcgta ccagccctc ctacgaaca cgctcgccga gctgtacgtg      1380
cgggagtaca tgcgggagca ggaccgcaag ccccggaatg ccacgccgc gccactgcgg      1440
gaggcgccca gcgccaacgc gtcctgggag cgcacaaaga ccacctctc gatcgagttc      1500
gcccggctgc agtttacgta taaccacata cagcgccacg tgaatgacat gctggggcgc      1560
atcgccgtcg cgtggtgcga gctgcagaac cagcagctga ctctctggaa cgaggccgc      1620
aagctcaacc ccaacgcat cgctccgcc accgtcgcc gccgggtgag cgcgcgcag      1680
ctcgagacg tcatggcgt ctccacgtgc gtgcccgtcg ccccgacaa cgtgatcgtg      1740
cagaactcga tgcgcgtcag ctcgggcgc gggacgtgct acagccgcc cctggtcagc      1800
tttcggtacg aagaccagg cccgtgatc gaggggcagc tgggcgagaa caacgagctg      1860
cgctcaccg gcgacgcgt cgagccgtgc accgtgggc accggcgcta ctcatcttc      1920
ggcgggggt acgtgtactt cgaggagtac gcgtactct accagctgag tgcggccgac      1980
gtcaccaccg tcagcacct catcgacct aacatcacca tgctggagga ccacgagttt      2040
gtgcccctg aggtctacac gcgccacgag atcaaggaca gggccctgct ggactacag      2100
gagggtccagc gccgcaacca gctgcacgac ctgcgtttg ccgacatcga cacggtcac      2160
cgcgccgacg ccaacgccgc catgttcgc gggctgtgc cgttcttcga ggggatgggg      2220
gacttggggc gcgcggtcgg caaggtagtc atgggagtag tggggggcgt ggtgtcgcc      2280
gtctcggggc tgtctcctt tatgtccaac cccttcgggg cgcttgccgt ggggtgtgtg      2340
gtcctggccg gcctggtcgc ggccttcttc gccttcgct acgtcctgca actgcaacgc      2400
aatcccatga aggccctgta tccgctcacc accaaggaac tcaagaattc cgaccccggg      2460
ggcgtggggc gggaggggga ggaaggcgc gaggggggc ggtttgacga ggccaagtgt      2520
gccgaggccc gagaaatgat ccgatatatg gctttggtgt cggccatgga gcgcacggaa      2580
cacaaggcca gaaagaagg cacgagcgcc ctgctcagct ccaaggtcac caacatggtt      2640
ctgcgcaagc gcaacaaagc caggtactct ccgtccaca acgaggacga ggccggagac      2700
gaagacgagc tctaa      2715

```

<210> 75
 <211> 904
 <212> PRT
 <213> Virus

<400> 75

ES 2 662 030 T3

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

ES 2 662 030 T3

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

Gln Pro Gln Tyr Tyr Leu Ala Thr Gly Gly Phe Leu Ile Ala Tyr Gln
435 440 445

Pro Leu Leu Ser Asn Thr Leu Ala Glu Leu Tyr Val Arg Glu Tyr Met
450 455 460

Arg Glu Gln Asp Arg Lys Pro Arg Asn Ala Thr Pro Ala Pro Leu Arg
465 470 475 480

Glu Ala Pro Ser Ala Asn Ala Ser Val Glu Arg Ile Lys Thr Thr Ser
485 490 495

Ser Ile Glu Phe Ala Arg Leu Gln Phe Thr Tyr Asn His Ile Gln Arg
500 505 510

His Val Asn Asp Met Leu Gly Arg Ile Ala Val Ala Trp Cys Glu Leu
515 520 525

Gln Asn His Glu Leu Thr Leu Trp Asn Glu Ala Arg Lys Leu Asn Pro
530 535 540

Asn Ala Ile Ala Ser Ala Thr Val Gly Arg Arg Val Ser Ala Arg Met
545 550 555 560

Leu Gly Asp Val Met Ala Val Ser Thr Cys Val Pro Val Ala Pro Asp
565 570 575

Asn Val Ile Val Gln Asn Ser Met Arg Val Ser Ser Arg Pro Gly Thr
580 585 590

Cys Tyr Ser Arg Pro Leu Val Ser Phe Arg Tyr Glu Asp Gln Gly Pro
595 600 605

Leu Ile Glu Gly Gln Leu Gly Glu Asn Asn Glu Leu Arg Leu Thr Arg
610 615 620

Asp Ala Leu Glu Pro Cys Thr Val Gly His Arg Arg Tyr Phe Ile Phe
625 630 635 640

Gly Gly Gly Tyr Val Tyr Phe Glu Glu Tyr Ala Tyr Ser His Gln Leu
645 650 655

Ser Arg Ala Asp Val Thr Thr Val Ser Thr Phe Ile Asp Leu Asn Ile
660 665 670

ES 2 662 030 T3

Thr Met Leu Glu Asp His Glu Phe Val Pro Leu Glu Val Tyr Thr Arg
 675 680 685
 His Glu Ile Lys Asp Ser Gly Leu Leu Asp Tyr Thr Glu Val Gln Arg
 690 695 700
 Arg Asn Gln Leu His Asp Leu Arg Phe Ala Asp Ile Asp Thr Val Ile
 705 710 715 720
 Arg Ala Asp Ala Asn Ala Ala Met Phe Ala Gly Leu Cys Ala Phe Phe
 725 730 735
 Glu Gly Met Gly Asp Leu Gly Arg Ala Val Gly Lys Val Val Met Gly
 740 745 750
 Val Val Gly Gly Val Val Ser Ala Val Ser Gly Val Ser Ser Phe Met
 755 760 765
 Ser Asn Pro Phe Gly Ala Leu Ala Val Gly Leu Leu Val Leu Ala Gly
 770 775 780
 Leu Val Ala Ala Phe Phe Ala Phe Arg Tyr Val Leu Gln Leu Gln Arg
 785 790 795 800
 Asn Pro Met Lys Ala Leu Tyr Pro Leu Thr Thr Lys Glu Leu Lys Thr
 805 810 815
 Ser Asp Pro Gly Gly Val Gly Gly Glu Gly Glu Glu Gly Ala Glu Gly
 820 825 830
 Gly Gly Phe Asp Glu Ala Lys Leu Ala Glu Ala Arg Glu Met Ile Arg
 835 840 845
 Tyr Met Ala Leu Val Ser Ala Met Glu Arg Thr Glu His Lys Ala Arg
 850 855 860
 Lys Lys Gly Thr Ser Ala Leu Leu Ser Ser Lys Val Thr Asn Met Val
 865 870 875 880
 Leu Arg Lys Arg Asn Lys Ala Arg Tyr Ser Pro Leu His Asn Glu Asp
 885 890 895
 Glu Ala Gly Asp Glu Asp Glu Leu
 900

<210> 76
 <211> 2715
 <212> ADN
 <213> Virus

<400> 76

ES 2 662 030 T3

atgcgcggag gaggactgat ctgcgctctg gtcgtgggag ctctggtggc tgctgtggct	60
tcggctgctc ccgctgctcc cgctgctccc cgcgcttcgg gaggagtggc tgctacggtc	120
gctgctaacg gaggacccgc ttgcgcgacc ccccccgtcc cctcgcccgc tacgacgaag	180
gctcgaaagc gaaaaacgaa aaagccccc aagcgaccg aagctacgcc ccccccgac	240
gctaacgcta cggtcgctgc tggacacgct acgctgcgcg ctcacctgcg agaaatcaag	300
gtcgaaaacg ctgatgctca gttctacgtg tgccccccc ccacgggagc tacggtgggtg	360
cagttcgaac agccccgcgc ctgccccacg cgccccgaag gacagaacta cacggaagga	420
atcgctgtgg tctttaagga aaacatcgct cctacaaat ttaaggctac gatgtactac	480
aaagacgtga cgggtgcgca ggtgtggttt ggacaccgct actcgagtt catgggaatc	540
tttgaagacc gcgctcccg tccctttgaa gaagtgatcg acaagatcaa cacgaaggga	600
gtctgccgct cgacggctaa gtacgtgcga aacaacatgg aaacgacggc tttccaccga	660
gacgaccacg aaacggacat ggaactgaag ccgcctaagg tcgctacgcg cacgtcgcga	720
ggatggcaca cgacggacct gaagtacaac cctcgcgag tggaagcttt tcatcgatac	780
ggaacgacgg tcaactgcat cgtcgaagaa gtggacgctc gatcggtgta cccctacgat	840
gaattcgtgc tggctacggg agacttcgtg tacatgtcgc ccttctacgg ataccgagaa	900
ggatcgcaac cggaacacac gtcgtacgct gctgaccgct ttaagcaggt cgacggattt	960
tacgctcgcg acctgacgac gaaggctcga gctacgtcgc ccacgacgcg caacctgctg	1020
acgacgcca agttcacggt ggcttgggac tgggtgcca agcgaccgcg tgtctgcacg	1080
atgacgaagt ggcaggaagt ggacgaaatg ctgcgcgctg aatacgagg atcgtttcgc	1140
ttttcgtcgg acgctatctc gacgacgttt acgacgaacc tgacggaata ctcgctgtcg	1200
cgcgtcgacc tgggagactg catcgacga gatgctcgcg aagctatcga ccgcatgttc	1260
gctcgcaagt acaacgctac gcacatcaag gtgggacagc ccagtacta cctggctacg	1320
ggaggatttc tgatcgctta ccagcccctg ctgtcgaaca cgctggctga actgtacgtg	1380
cgagaataca tgcgagaaca ggaccgcaag ccccgaaatg ctacgcccgc tcccctgcga	1440
gaagctccct cggctaacgc ttcggtggaa cgcacatcaaga cgacgtcgtc gatcgaaattt	1500
gctcgactgc agttcacgta taaccacatc cagcgccacg tgaatgacat gctgggacgc	1560
atcgctgtcg cttggtgcga actgcagaac cacgaactga cgctgtggaa cgaagctcgc	1620
aagctgaacc ccaacgctat cgcttcgggt acggtcggac gacgagtgtc ggctcgcgtg	1680

ES 2 662 030 T3

ctgggagacg tcatggctgt ctcgacgtgc gtgcccgtcg ctcccagaaa cgtgatcgtg	1740
cagaactcga tgcgcgtctc gtcgcgaccc ggaacgtgct actcgcgccc cctgggtctcg	1800
ttccgatacg aagaccaggg acccctgacg gaaggacagc tgggagaaaa caacgaactg	1860
cgcctgacgc gcgacgtctc ggaaccctgc acggtgggac accgacgcta ctttatcttt	1920
ggaggaggat acgtgtactt tgaagaatac gcttactcgc accagctgtc gcgcgctgac	1980
gtcacgacgg tctcgacgtt tatcgacctg aacatcacga tgctggaaga ccacgaattc	2040
gtgcccctgg aagtctacac gcgccacgaa atcaaggact cgggactgct ggactacacg	2100
gaagtccagc gccgcaacca gctgcacgac ctgcgcttcg ctgacatcga cacggtcatc	2160
cgcgctgacg ctaacgctgc tatgtttgct ggactgtgcg ctttttttga aggaatggga	2220
gacctgggac gcgctgtcgg aaaggtcgtc atgggagtcg tgggaggagt ggtgtcggct	2280
gtctcgggag tgtcgtcgtt catgtcgaac ccctttggag ctctggctgt gggactgctg	2340
gtcctggctg gactggctgc tgcttttttt gcttttcgct acgtcctgca actgcaacgc	2400
aatcccatga aggtctctga tcccctgacg acgaaggaa tgaagacgtc ggaccccgga	2460
ggagtgggag gagaaggaga agaaggagct gaaggaggag gattcgacga agctaagctg	2520
gctgaagctc gagaaatgat ccgatatatg gctctggtgt cggctatgga acgcacggaa	2580
cacaaggcta gaaagaaggg aacgtcggct ctgctgtcgt cgaaggtcac gaacatggtc	2640
ctgcgcaagc gcaacaaagc tagatactcg cccctgcaca acgaagacga agctggagac	2700
gaagacgaac tgtaa	2715

<210> 77
 <211> 1182
 <212> ADN
 <213> Virus

<400> 77

atggggcggt tgacctccgg cgtcgggacg gcggccctgc tagttgtcgc ggtgggactc	60
cgcgtcgtct gcgccaaata cgccttagca gaccctcgc ttaagatggc cgatcccaat	120
cgatttcgcy ggaagaacct tccggttttg gaccagctga ccgaccccc cggggtgaag	180
cgtgtttacc acattcagcc gagcctggag gaccggttc agccccccag catcccgatc	240
aotgtgtact aogcagtget ggaaogtgcc tgccgcagcg tgetcctaca tgcccatcg	300
gagggccccc agatcgtcgy cggggcttcg gacgaggccc gaaagcacac gtacaacctg	360
accatcgctt ggtatcgcat gggagacaat tgcgctatcc ccatcacggt tatggaatac	420
accgagtgcc cctacaacaa gtcgttgggg gtctgcccc tccgaacgca gccccgctgg	480
agctactatg acagcttttag cgcgctcagc gaggataacc tgggattcct gatgcacgcc	540

ccgccttcg agaccgcggtg tacgtacctg cggctagtga agataaacga ctggacggag 600
atcacacaat ttatcctgga gcaccgggcc cgcgcctcct gcaagtacgc tctccccctg 660
cgcatccccc cggcagcgtg cctcacctcg aaggcctacc aacagggcgt gacggtcgac 720
agcatcggga tgctaccccg ctttatcccc gaaaaccagc gcaccgtcgc cctatacagc 780
ttaaaaatcg ccgggtggca cggccccaag cccccgtaca ccagcaccct gctgccgccg 840
gagctgtcgc acaccaccaa cgccacgcaa cccgaactcg ttccggaaga ccccgaggac 900
tcggccctct tagaggatcc cgcggggacg gtgtcttcgc agatcccccc aaactggcac 960
atcccgtcga tccaggacgt cgcgccgcac cagcccccag cgcggggcag caaccggggc 1020
ctgatcatcg gcgcgtggc cggcagtcac ctggcggtgc tggtcacgag cgggtattgcg 1080
ttttgggtac gccgccgcgc tcagatggcc cccaagcgcc tacgtctccc ccacatccgg 1140
gatgacgacg cgccccctc gcaccagcca ttgttttact ag 1182

<210> 78
<211> 393
<212> PRT
<213> Virus

<400> 78

Met Gly Arg Leu Thr Ser Gly Val Gly Thr Ala Ala Leu Leu Val Val
1 5 10 15
Ala Val Gly Leu Arg Val Val Cys Ala Lys Tyr Ala Leu Ala Asp Pro
20 25 30
Ser Leu Lys Met Ala Asp Pro Asn Arg Phe Arg Gly Lys Asn Leu Pro
35 40 45
Val Leu Asp Gln Leu Thr Asp Pro Pro Gly Val Lys Arg Val Tyr His
50 55 60
Ile Gln Pro Ser Leu Glu Asp Pro Phe Gln Pro Pro Ser Ile Pro Ile
65 70 75 80
Thr Val Tyr Tyr Ala Val Leu Glu Arg Ala Cys Arg Ser Val Leu Leu
85 90 95
His Ala Pro Ser Glu Ala Pro Gln Ile Val Arg Gly Ala Ser Asp Glu
100 105 110
Ala Arg Lys His Thr Tyr Asn Leu Thr Ile Ala Trp Tyr Arg Met Gly
115 120 125

ES 2 662 030 T3

Asp Asn Cys Ala Ile Pro Ile Thr Val Met Glu Tyr Thr Glu Cys Pro
 130 135 140
 Tyr Asn Lys Ser Leu Gly Val Cys Pro Ile Arg Thr Gln Pro Arg Trp
 145 150 155 160
 Ser Tyr Tyr Asp Ser Phe Ser Ala Val Ser Glu Asp Asn Leu Gly Phe
 165 170 175
 Leu Met His Ala Pro Ala Phe Glu Thr Ala Gly Thr Tyr Leu Arg Leu
 180 185 190
 Val Lys Ile Asn Asp Trp Thr Glu Ile Thr Gln Phe Ile Leu Glu His
 195 200 205
 Arg Ala Arg Ala Ser Cys Lys Tyr Ala Leu Pro Leu Arg Ile Pro Pro
 210 215 220
 Ala Ala Cys Leu Thr Ser Lys Ala Tyr Gln Gln Gly Val Thr Val Asp
 225 230 235 240
 Ser Ile Gly Met Leu Pro Arg Phe Ile Pro Glu Asn Gln Arg Thr Val
 245 250 255
 Ala Leu Tyr Ser Leu Lys Ile Ala Gly Trp His Gly Pro Lys Pro Pro
 260 265 270
 Tyr Thr Ser Thr Leu Leu Pro Pro Glu Leu Ser Asp Thr Thr Asn Ala
 275 280 285
 Thr Gln Pro Glu Leu Val Pro Glu Asp Pro Glu Asp Ser Ala Leu Leu
 290 295 300
 Glu Asp Pro Ala Gly Thr Val Ser Ser Gln Ile Pro Pro Asn Trp His
 305 310 315 320
 Ile Pro Ser Ile Gln Asp Val Ala Pro His His Ala Pro Ala Ala Pro
 325 330 335
 Ser Asn Pro Gly Leu Ile Ile Gly Ala Leu Ala Gly Ser Thr Leu Ala
 340 345 350
 Val Leu Val Ile Gly Gly Ile Ala Phe Trp Val Arg Arg Arg Ala Gln
 355 360 365
 Met Ala Pro Lys Arg Leu Arg Leu Pro His Ile Arg Asp Asp Asp Ala
 370 375 380
 Pro Pro Ser His Gln Pro Leu Phe Tyr
 385 390

<210> 79
<211> 1182
<212> ADN
<213> Virus

5

<400> 79

```

atgggacgtc tgacgtcggg agtcggaacg gctgctctgc tggtcgtcgc tgtgggactg      60
cgcgctcgtc gcgctaaata cgctctggct gacccctcgc tgaagatggc tgatcccaat      120
cgattccgcg gaaagaacct gcccgctcctg gaccagctga cggacccccc cggagtgaa      180
cgtgtctacc acatccagcc ctgctggaa gaccccttc agccccctc gatccccatc      240
acgggtgtact acgctgtgct ggaacgtgct tgccgctcgg tgctgtgca tgcctcctcg      300
gaagctcccc agatcgtcgc cggagcttcg gacgaagctc gaaagcacac gtacaacctg      360
acgatcgctt ggtatcgcat gggagacaat tgcgctatcc ccatcacggt catggaatac      420
acggaatgcc cctacaacaa gtcgctggga gtctgcccc tccgaacgca gccccgctgg      480
tcgtactatg actcgttctc ggctgtctcg gaagataacc tgggatttct gatgcacgct      540
cccgttttg aaacggctgg aacgtacctg cgactgggta agatcaacga ctggacggaa      600
atcacgcaat tcctcctgga acaccgagct cgcgcttcgt gcaagtagc tctgccccctg      660
cgcaccccc cgcgtgcttg cctgacgtcg aaggcttacc aacagggagt gacggtcgac      720
tcgatcggaa tgctgcccc ctcatcccc gaaaaccagc gcacggtcgc tctgtactcg      780
ctgaaaatcg ctggatggca cggacccaag cccccctaca cgtcgacgct gctgcccccc      840
gaactgtcgg acacgacgaa cgctacgcaa cccgaactgg tccccgaaga cccgaagac      900
tcggctctgc tggaagatcc cgctggaacg gtgtcgtcgc agatcccccc caactggcac      960
atccccctga tccaggacgt cgctccccac cacgctcccg ctgctccctc gaaccccgga      1020
ctgatcatcg gagctctggc tggatcgacg ctggctgtgc tggatcatcg aggaatcgct      1080
ttctgggtcc gccgcgcgc tcagatggct cccaagcgcc tgcgtctgcc ccacatccga      1140
gatgacgacg ctccccctc gcaccagccc ctgttctact ag                          1182

```

10

<210> 80
<211> 387
<212> ADN
<213> Virus del papiloma humano tipo 16

15

<400> 80

```

ggtagcgccg ccaccatgga gacagacaca ctctgctat gggtagctg gctctgggtt      60
ccaggttcca ctggtgacgg atccatgcat ggagatacac ctacattgca tgaatatatg      120
ttagatttgc aaccagagac aactgatctc tactgttatg agcaattaaa tgacagctca      180
gaggaggagg atgaaataga tggccagct ggacaagcag aaccggacag agcccattac      240
aatattgtaa ctttttggtg caagtgtgac tctacgcttc ggttgtgctg acaaagcaca      300
cacgtagaca ttctactttt ggaagacctg ttaatgggca cactaggaat tgtgtgcccc      360
atctgctctc agaagcccta agaattc                                387

```

ES 2 662 030 T3

<210> 81
 <211> 387
 <212> ADN
 <213> Secuencia artificial

5

<220>
 <223> HPV-16 E7 O1

10

<400> 81

ggtagcgccg ccaccatgga aacggacacg ctgctgctgt gggtcctgct gctgtggggtc	60
cccggatcga cgggagacgg atcgatgcat ggagacacgc ccacgctgca tgaatacatg	120
ctggacctgc aaccgaaac gacggacctg tactgctacg aacaactgaa cgactcgctcg	180
gaagaagaag acgaaatcga cggacccgct ggacaagctg aaccgacag agctcattac	240
aacatcgta cgttctgctg caagtgcgac tcgacgctgc gactgtgctg ccaatcgacg	300
cacgtcgaca tccgtacgct ggaagacctg ctgatgggaa cgctgggaat cgtgtgcccc	360
atctgctcgc agaagcccta agaattc	387

<210> 82
 <211> 387
 <212> ADN
 <213> Secuencia artificial

15

<220>
 <223> VPH16 E7 02

20

<400> 82

ggtagcgccg ccaccatgga aacggacacg ctgctgctgt gggtcctgct gctgtggggtc	60
cccggatcga cgggagacgg atcgatgcat ggagatacgc ctacgctgca tgaatatatg	120
ctggatctgc aaccgaaac gacggatctg tactgttatg aacaactgaa tgactcgctcg	180
gaagaagaag atgaaatcga tggacccgct ggacaagctg aaccgacag agctcattac	240
aatatcgta cgttttgttg caagtgtgac tcgacgctgc gactgtgctg ccaatcgacg	300
cacgtcgaca tccgtacgct ggaagacctg ctgatgggaa cgctgggaat cgtgtgcccc	360
atctgctcgc agaagcccta agaattc	387

25

<210> 83
 <211> 417
 <212> ADN
 <213> Secuencia artificial

30

<220>
 <223> VPH-16 E7 03

35

<400> 83

ES 2 662 030 T3

```

gggtaccgccg ccaccatgga gacggacacg ctectgctct ggggtactgct gctctggggtt      60
cctggatcga cgggattgtg gacggatcga tgcattggaga tacgcctacg ctccatgaat      120
atatgctcga tctccaacct ggttgagacg acggatctct actgttatga gcaactcaat      180
gactcgctcg aggaggagga tgaattcata gatggacctg ctggacaagc agaacctgac      240
agagccatt acaatattgt aacgtttgag aattgttgca agtgtgactc gacgctccgg      300
ctctgcgtac aatcgacgca cgtagacatt cgtccctcta cgctcgaaga cctgctcatg      360
ggaacgctcg gaattgtgtg ccccatctgc tcgcagaagt gtgcccccta agaattc      417

```

<210> 84
 <211> 387
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> VPH-16 E7 W

<400> 84

```

gggtaccgccg ccaccatgga gactgatact ttattattat ggggtattatt attatggggtt      60
ccaggtagta ctggtgatgg cagtatgcat ggcgatactc caactttaca tgagtatatg      120
ttagatttac aaccagagac tactgattta tattgttatg agcaattaaa tgatagcagt      180
gaggaggagg atgagataga tgggccagcg ggccaagcag agccggatcg ggcgcattat      240
aatatagtaa ctttctgttg taagtgtgat agtactttac gggtatgtgt acaaagcact      300
cacgtagata tacggacttt agaggattta ttaatgggca ctttaggcatt agtatgtcca      360
atatgtagtc agaagccata agaattc      387

```

<210> 85
 <211> 1182
 <212> ADN
 <213> Virus del herpes simple tipo 2

<400> 85

```

atggggcggtt tgacctccgg cgtcgggacg gcggccctgc tagttgtcgc ggtgggactc      60
cgcgtcgtct gcgccaaata cgccttagca gaccctcgc ttaagatggc cgatcccaat      120

```

cgatttcgcg ggaagaacct tccggttttg gaccagctga ccgaccccc cggttggaag 180
 cgtgtttacc acattcagcc gagcctggag gaccggttcc agccccccag catcccgatc 240
 actgtgtact acgcagtgtc ggaacgtgcc tgccgcagcg tgctcctaca tgccccatcg 300
 gagggccccc agatcgtgcg cggggcttcg gacgaggccc gaaagcacac gtacaacctg 360
 accatcgctt ggtatcgcat gggagacaat tgcgctatcc ccatcacggt tatggaatac 420
 accgagtgcc cctacaacaa gtcgttgggg gtctgcccc tccgaacgca gccccgctgg 480
 agctactatg acagcttttag cgccgtcagc gaggataacc tgggattcct gatgcacgcc 540
 cccgccttcg agaccgctgg tacgtacctg cggctagtga agataaacga ctggacggag 600
 atcacacaat ttatcctgga gcaccgggccc cgccgctcct gcaagtacgc tctccccctg 660
 cgcacccccc cggcagcgtg cctcacctcg aaggcctacc aacaggggcg gacggtcgac 720
 agcatcgcca tgctaccccc ctttatcccc gaaaaccagc gcaccgtcgc cctatacagc 780
 ttaaaaatcg ccgggtggca cggccccaag cccccgtaca ccagcaccct gctgccgccg 840
 gagctgtccg acaccaccaa cggcacgcaa cccgaactcg ttccggaaga ccccgaggac 900
 tcggccctct tagaggatcc cgccgggacg gtgtcttcgc agatccccc aaactggcac 960
 atcccgctga tccaggacgt cgccgcgcac cagccccccg ccgccccag caaccgggc 1020
 ctgatcatcg gcgcgctggc cggcagtacc ctggcgggtg tggtcatcgg cggtatcgcg 1080
 ttttgggtac gccgcgcgc tcagatggcc cccaagcgcc tacgtctccc ccacatccgg 1140
 gatgacgacg cggccccctc gcaccagcca ttgttttact ag 1182

<210> 86
 <211> 1182
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> VHS-2 gD2 O1

<400> 86

atgggacgtc tgacgtcggg agtcggaacg gctgctctgc tggtcgtcgc tgtgggactc 60
 cgcgtcgtct gcgctaaata cgctctggct gaccctcgc tgaagatggc tgaccccaac 120
 cgatttcgcg gaaagaacct gccgctctg gaccagctga cggaccccc cggagtgaag 180
 cgtgtctacc acatccagcc ctgctggaa gacccttctc agccccctc gatccccatc 240
 acgggtgtact acgctgtgct ggaacgtgct tgccgctcgg tgctcctcca tgctccctcg 300
 gaagctcccc agatcgtgcg cggagcttcg gacgaagctc gaaagcacac gtacaacctg 360
 acgatcgctt ggtaccgcat gggagacaac tgcgctatcc ccatcacggt catggaatac 420
 acggaatgcc cctacaacaa gtcgctcgga gtctgcccc tccgaacgca gccccgctgg 480

tcgtactacg actcgttttc ggctgtctcg gaagacaacc tgggatttct gatgcacgct	540
cccgtttttg aaacggctgg aacgtacctg cgactcgtga agatcaacga ctggacggaa	600
atcacgcaat ttatcctgga acaccgagct cgcgcttcgt gcaagtacgc tctccccctg	660
cgcacccccc ccgctgcttg cctcacgtcg aaggcttacc aacagggagt gacggtcgac	720
tcgatcggaa tgctcccccg ctttatcccc gaaaaccagc gcacggtcgc tctctactcg	780
ctcaaaatcg ctggatggca cggacccaag cccccctaca cgtcgacgct gctgcccccc	840
gaactgtcgg acacgacgaa cgctacgcaa cccgaactcg tccccgaaga ccccgaagac	900
tcggctctcc tcgaagacc cgcgtggaacg gtgtcgtcgc agatcccccc caactggcac	960
atccccctga tccaggacgt cgcgtccccc cagcgtcccg ctgctccctc gaaccccgga	1020
ctgatcatcg gagctctggc tggatcgacg ctggctgtgc tggtcacgag aggaatcgct	1080
ttttgggtcc gccgcgcgc tcagatggct cccaagcgcc tccgtctccc ccacatccga	1140
gacgacgacg ctccccctc gcaccagccc ctcttttact ag	1182

<210> 87
 <211> 1182
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> VHS-2 gD2 02

<400> 87

atgggacgtc tgacgtcggg agtcggaacg gctgctctgc tggtcgtcgc tgtgggactg	60
cgcgtcgtct gcgttaaata cgtctctggct gaccctctgc tgaagatggc tgatcccaat	120
cgatttcgcg gaaagaacct gcccgctctg gaccagctga cggaccccc cggagtgaag	180
cgtgtctacc acatccagcc ctgctggaa gaccctttc agccccctc gatccccatc	240
acgggtgtact acgtgtgtc ggaacgtgtc tgccgctcgg tgetgtgca tgetccctcg	300
gaagctcccc agatcgtcgc cggagcttcg gacgaagctc gaaagcacac gtacaacctg	360
acgatcgctt ggtatcgcat gggagacaat tgcgctatcc ccatcacggt catggaatac	420
acggaatgcc cctacaacaa gtcgctggga gtctgcccc tccgaacgca gcccgctgg	480
tcgtactatg actcgttttc ggctgtctcg gaagataacc tgggatttct gatgcacgct	540
cccgtttttg aaacggctgg aacgtacctg cgactcgtga agatcaacga ctggacggaa	600
atcacgcaat ttatcctgga acaccgagct cgcgcttcgt gcaagtacgc tctgccccctg	660
cgcacccccc ccgctgcttg cctgacgtcg aaggcttacc aacagggagt gacggtcgac	720
tcgatcggaa tgctgccccg ctttatcccc gaaaaccagc gcacggtcgc tctgtactcg	780

ES 2 662 030 T3

```

ctgaaaatcg ctggatggca cggacccaag cccccctaca cgtcgacgct gctgcccccc 840
gaactgtcgg acacgacgaa cgctacgcaa cccgaactgg tccccgaaga ccccgaaagac 900
tcggctctgc tggaagatcc cgctggaacg gtgtcgtcgc agatcccccc caactggcac 960
atccccctga tccaggacgt cgctccccac cagctccccg ctgctccctc gaaccccgga 1020
ctgatcatcg gagctctggc tggatcgacg ctggctgtgc tggtcacgag aggaatcgct 1080
ttttgggtcc gccgccgcgc tcagatggct cccaagcgcc tgcgtctgcc ccacatccga 1140
gatgacgacg cccccccctc gcaccagccc ctgttttact ag 1182

```

<210> 88
 <211> 1182
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> VHS-2 gD2 03

<400> 88

```

atgggacgtc tcacgtcggg agtcggaacg gcggccctgc tcgttgctgc ggtgggactc 60
cgcgtcgtct gcgccaata cgccctcgca gaccctcgc tcaagatggc cgatcccaat 120
cgatttcgag gaaagaacct cctgttctc gaccagctga cggaccccc cggagtgaag 180
cgtgtttacc acattcagcc ttcgtggag gacccttcc agccccctc gatccctatc 240
acgggtgtact acgcagtgtc ggaacgtgcc tgcgcctcgc tgctcctcca tgcccttcg 300
gaggcccccc agatcgtgcg cggagcttcg gacgaggccc gaaagcacac gtacaacctg 360
acgatcgctt ggtatcgcat gggagacaat tgcgtatcc ccatcacggt tatggaatac 420
acggagtgcc cctacaacaa gtcgctcgga gtctgcccc tccgaacgca gccccgctgg 480
tcgtactatg actcgttttc ggccgtctcg gaggataacc tgggattcct gatgcacgcc 540
ccgccttcg agacggcggg aacgtacctg cggctcgtga agataaacga ctggacggag 600
atcacgcaat ttatcctgga gcaccgggcc cgcgcctcgt gcaagtagc tctccccctg 660
cgcatecccc ctgcagcgtg cctcacgtcg aaggcctacc aacagggagt gacggtcgac 720
tcgatcgga tgcctccccg ctttatcccc gaaaaccagc gcacggtcgc cctctactcg 780
ctcaaaatcg ccgatggca cggacccaag ccccttaca cgtcgacgct gctgcctcct 840
gagctgtcgg acacgacgaa cgccacgcaa cccgaactcg ttctgaaga ccccgaggac 900
tcggccctcc tagaggatcc cgccggaacg gtgtcgtcgc agatcccccc taactggcac 960
atccccctga tccaggacgt cgccctcac cagcccccg ccgccccctc gaacctgga 1020
ctgatcatcg gagcgtggc cggatcgacg ctggcgggtg tggtcacgag aggaattgag 1080
ttttgggtac gccgccgcgc tcagatggcc cccaagcgcc tccgtctccc ccacatccgg 1140
gatgacgacg cccccccctc gcaccagcct ctcttttact ag 1182

```

<210> 89
 <211> 1182
 <212> ADN
 <213> Secuencia artificial

<220>
<223> VHS-2 gD2 W

<400> 89

5

```

atggggcggt tgactagtgg cgtagggact gggcggttat tagtagtagc ggtaggctta      60
cgggtagtat gtgcaaaata tgcgttagca gatccaagtt taaagatggc ggatccaaat      120
cggttccggg ggaagaattt accgggtattg gatcagttaa ctgatccacc aggggtaaag      180
cgggtatatc acatacagcc gagcttagag gatccgttcc agccaccaag cataccgata      240
actgtatatc atgcagtatt agagcgggcg tgcgggagcg tattattaca tgcaccaagt      300
gaggcgccac agatagtacg gggggcaagt gatgaggcgc ggaagcacac ttataattta      360
actatagcat ggtatcggat gggcgataat tgtgcgatac caataactgt aatggagtat      420
actgagtgtc catataataa gagtttgggg gtatgtccaa tacggactca gccacggtgg      480
agctattatg atagcttcag cgcagtaagc gaggataatt taggcttctt aatgcacgcg      540
ccagcattcg agactgcggg tacttattta cgggttagtaa agataaatga ttggactgag      600
ataactcaat tcatattaga gcaccgggca cgggcgagtt gtaagtatgc attaccatta      660
cggataccac cggcagcgtg ttaactagt aaggcatatc aacagggcgt aactgtagat      720
agcatagggg tgttaccacg gttcatacca gagaatcagc ggactgtagc gttatatagc      780
ttaaaaatag cagggtggca cggcccaaag ccaccgtata ctagcacttt attaccgccg      840
gagttaagtg atactactaa tgcgactcaa ccagagttag taccggagga tccagaggat      900
agtgcattat tagaggatcc agcggggact gtaagtagtc agataccacc aaattggcac      960
ataccgagta tacaggatgt agcgccgcac caccgaccag cggcaccaag caatccgggc     1020
ttaataatag gcgcgttagc aggcagtact ttagcggtat tagtaatagg cggtatagcg     1080
ttctgggtac ggcggcgggc gcagatggcg ccaagcggtt tacgggttacc acacatacgg     1140
gatgatgatg cgccaccaag tcaccagcca ttgttctatt ag                          1182

```

<210> 90
<211> 41
<212> ADN
<213> Secuencia artificial

10

<220>
<223> Cebador directo común

15

<400> 90
ttgaataggt accgcgcca ccatggagac cgacaccctc c 41

<210> 91
<211> 24
<212> ADN
<213> Secuencia artificial

20

<220>
<223> ODN-7909

25

<400> 91
tcgtcgtttt gtcgtttgt cggt 24

REIVINDICACIONES

1. Un sistema de construcciones para determinar la eficiencia de la traducción o la preferencia fenotípica de diferentes codones sinónimos, comprendiendo el sistema una pluralidad de construcciones sintéticas, comprendiendo, cada una, una secuencia reguladora que está conectada operativamente a un polinucleótido indicador y dicha secuencia reguladora no está unida operativamente a una repetición en tándem de un codón fusionado en fase de lectura con dicho polinucleótido indicador, en el que el polinucleótido indicador de una primera construcción comprende una primera secuencia codificante para interrogar la eficiencia de la traducción o la preferencia fenotípica de un primer codón de interrogación que codifica un primer aminoácido, en el que el polinucleótido indicador de una segunda construcción comprende una segunda secuencia codificante para interrogar la eficiencia de la traducción o la preferencia fenotípica de un segundo codón de interrogación que codifica el primer aminoácido, en el que las secuencias codificantes primera y segunda codifican la misma secuencia de aminoácidos, que define una proteína indicadora, en el que la primera secuencia codificante comprende el primer codón de interrogación para codificar el primer aminoácido en una o más posiciones de la secuencia de aminoácidos, en el que la segunda secuencia codificante comprende el segundo codón de interrogación que codifica el primer aminoácido en una o más posiciones de la secuencia de aminoácidos y en el que las secuencias de codificación primera y segunda difieren entre sí solo en la elección del primer codón de interrogación o el segundo codón de interrogación para codificar el primer aminoácido en la posición o posiciones correspondientes en la secuencia de aminoácidos y en el que la primera secuencia codificante comprende el mismo número de primeros codones de interrogación que el número de segundos codones de interrogación en la segunda secuencia codificante.
2. Un sistema de acuerdo con la reivindicación 1, en el que el sistema de construcciones comprende una o más construcciones sintéticas para interrogar la eficiencia de la traducción o la preferencia fenotípica de uno o más codones de interrogación que codifican el primer aminoácido y/o un número correspondiente de construcciones sintéticas como el número de codones sinónimos que normalmente codifican el primer aminoácido.
3. Un sistema de acuerdo con la reivindicación 1, en el que al menos el 10 % de los codones que codifican el primer aminoácido en la secuencia codificante de construcciones sintéticas individuales son el mismo codón de interrogación.
4. Un sistema de acuerdo con la reivindicación 1, en el que la secuencia codificante de polinucleótidos indicadores individuales codifica un polipéptido que confiere un fenotipo a una célula o un tejido en los que se expresa la secuencia codificante.
5. Un sistema de acuerdo con la reivindicación 4, en el que el polipéptido se selecciona entre una proteína indicadora que, cuando está presente en una célula o tejido, es detectable ya sea por su presencia o actividad.
6. Un sistema de acuerdo con la reivindicación 5, en el que la proteína indicadora se selecciona entre una proteína indicadora quimioluminiscente tal como luciferasa, una proteína fluorescente tal como la proteína fluorescente verde, una proteína indicadora enzimática tales como cloranfenicol acetil transferasa, β -galactosidasa, fosfatasa alcalina placentaria secretada, β -lactamasa o un factor de crecimiento tales como la hormona de crecimiento humana.
7. Un sistema de acuerdo con la reivindicación 1, en el que la secuencia codificante de polinucleótidos indicadores individuales codifica un polipéptido que confiere un fenotipo a una célula o un tejido en los que no se expresa la secuencia codificante.
8. Un sistema de acuerdo con la reivindicación 7, en el que el polipéptido es un polipéptido asociado a fenotipo que es el objeto de la producción de un fenotipo seleccionado o un fenotipo de la misma clase que el fenotipo seleccionado.
9. Un método para determinar la eficiencia de la traducción de un primer codón con respecto a un segundo codón en una célula de interés, en el que el primer codón y el segundo codón codifican el mismo aminoácido, comprendiendo el método:
 - proporcionar una pluralidad de construcciones sintéticas, comprendiendo, cada una, una secuencia reguladora que está conectada operativamente a un polinucleótido indicador y dicha secuencia reguladora no está unida operativamente a una repetición en tándem de un codón fusionado en fase de lectura con dicho polinucleótido indicador, en el que el polinucleótido indicador de una primera construcción comprende una primera secuencia codificante para interrogar la eficiencia de la traducción del primer codón, en el que el polinucleótido indicador de una segunda construcción comprende una segunda secuencia codificante para interrogar la eficiencia de la traducción del segundo codón, en el que las secuencias codificantes primera y segunda codifican la misma secuencia de aminoácidos, que define una proteína indicadora, en el que la primera secuencia codificante comprende el primer codón para codificar el primer aminoácido en una o más posiciones de la secuencia de aminoácidos, en el que la segunda secuencia codificante comprende el segundo codón para codificar el primer aminoácido en una o más posiciones de la secuencia de aminoácidos y en el que las secuencias de codificación primera y segunda difieren entre sí en la elección del primer codón o el segundo codón para codificar el primer

aminoácido en la posición o posiciones correspondientes en la secuencia de aminoácidos y en el que la primera secuencia codificante comprende el mismo número de primeros codones de interrogación que el número de segundos codones de interrogación en la segunda secuencia codificante;

- introducir la primera construcción en una célula del mismo tipo que la célula de interés;
- introducir la segunda construcción en una célula del mismo tipo que la célula de interés;
- medir la expresión de la proteína indicadora de la primera construcción y de la segunda construcción en la célula; y
- determinar la eficiencia de la traducción del primer codón y la eficiencia de la traducción del segundo codón basándose en la expresión medida de la proteína indicadora en la célula, para determinar de este modo la eficiencia de la traducción del primer codón con respecto al segundo codón en la célula de interés.

10. Un método de acuerdo con la reivindicación 9, que comprende adicionalmente determinar una comparación de eficiencias de la traducción de codones sinónimos individuales en la célula de interés.

11. Un método para determinar la eficiencia de la traducción de un primer codón y un segundo codón en un primer tipo celular con respecto a un segundo tipo celular, comprendiendo el método:

- proporcionar una pluralidad de construcciones sintéticas, comprendiendo cada una una secuencia reguladora que está conectada operativamente a un polinucleótido indicador y dicha secuencia reguladora no está unida operativamente a una repetición en tándem de un codón fusionado en fase de lectura con dicho polinucleótido indicador, en el que el polinucleótido indicador de una primera construcción comprende una primera secuencia codificante para interrogar la eficiencia de la traducción del primer codón, en el que el polinucleótido indicador de una segunda construcción comprende una segunda secuencia codificante para interrogar la eficiencia de la traducción del segundo codón, en el que las secuencias codificantes primera y segunda codifican la misma secuencia de aminoácidos, que define una proteína indicadora, en el que la primera secuencia codificante comprende el primer codón para codificar el primer aminoácido en una o más posiciones de la secuencia de aminoácidos, en el que la segunda secuencia codificante comprende el segundo codón para codificar el primer aminoácido en una o más posiciones de la secuencia de aminoácidos y en el que las secuencias de codificación primera y segunda difieren entre sí en la elección del primer codón o el segundo codón para codificar el primer aminoácido en la posición o posiciones correspondientes en la secuencia de aminoácidos y en el que la primera secuencia codificante comprende el mismo número de primeros codones de interrogación que el número de segundos codones de interrogación en la segunda secuencia codificante;
- introducir por separado la primera construcción en el primer tipo celular y en el segundo tipo celular;
- introducir por separado la segunda construcción en el primer tipo celular y en el segundo tipo celular;
- medir la expresión de la proteína indicadora en el primer tipo celular y en el segundo tipo celular a los que se proporcionó la primera construcción;
- medir la expresión de la proteína indicadora en el primer tipo celular y en el segundo tipo celular a los que se proporcionó la segunda construcción;
- determinar la eficiencia de la traducción del primer codón en el primer tipo celular y en el segundo tipo celular basándose en la expresión medida de la proteína indicadora en el primer tipo celular y en el segundo tipo celular, respectivamente, a los que se proporcionó la primera construcción, para determinar de este modo la eficiencia de la traducción del primer codón en el primer tipo celular con respecto al segundo tipo celular; y
- determinar la eficiencia de la traducción del segundo codón en el primer tipo celular y en el segundo tipo celular basándose en la expresión medida de la proteína indicadora en el primer tipo celular y en el segundo tipo celular, respectivamente, a los que se proporcionó la segunda construcción, para determinar de este modo la eficiencia de la traducción del segundo codón en el primer tipo celular con respecto al segundo tipo celular.

12. Un método de acuerdo con la reivindicación 11, que comprende adicionalmente determinar una comparación de eficiencias de la traducción de codones sinónimos individuales en el primer tipo celular con respecto al segundo tipo celular.

13. Un método para determinar la preferencia de un primer codón con respecto a la preferencia de un segundo codón para producir un fenotipo seleccionado en un organismo de interés o parte del mismo, en el que el primer codón y el segundo codón codifican el mismo aminoácido, comprendiendo el método:

- proporcionar una pluralidad de construcciones sintéticas, comprendiendo cada una una secuencia reguladora que está conectada operativamente a un polinucleótido indicador y dicha secuencia reguladora no está unida operativamente a una repetición en tándem de un codón fusionado en fase de lectura con dicho polinucleótido indicador, en el que el polinucleótido indicador de una primera construcción comprende una primera secuencia codificante para interrogar la preferencia fenotípica del primer codón, en el que el polinucleótido indicador de una segunda construcción comprende una segunda secuencia codificante para interrogar la preferencia fenotípica del segundo codón, en el que las secuencias de codificación primera y segunda codifican la misma secuencia de aminoácidos, que define una proteína indicadora, que produce o que se prevé que produce, el fenotipo seleccionado o un fenotipo de la misma clase que el fenotipo seleccionado, en el que la primera secuencia codificante comprende el primer codón para codificar el primer aminoácido en una o más posiciones de la secuencia de aminoácidos, en el que la segunda secuencia codificante comprende el segundo codón para

codificar el primer aminoácido en una o más posiciones de la secuencia de aminoácidos y en el que las secuencias de codificación primera y segunda difieren entre sí en la elección del primer codón o el segundo codón para codificar el primer aminoácido en la posición o posiciones correspondientes en la secuencia de aminoácidos y en el que la primera secuencia codificante comprende el mismo número de primeros codones de interrogación que el número de segundos codones de interrogación en la segunda secuencia codificante;

- introducir la primera construcción en un primer organismo de ensayo o parte del mismo, en el que el organismo de ensayo se selecciona entre el grupo que consiste en un organismo de la misma especie que el organismo de interés y un organismo que está relacionado con el organismo de interés;

- introducir la segunda construcción en un segundo organismo de ensayo o parte del mismo, en el que el segundo organismo de ensayo es del mismo tipo que el primer organismo;

- determinar la cualidad del fenotipo correspondiente presentada por el primer organismo de ensayo o parte del mismo y por el segundo organismo de ensayo o parte del mismo; y

- determinar la preferencia fenotípica del primer codón y la preferencia fenotípica del segundo codón basándose, respectivamente, en la cualidad del fenotipo correspondiente mostrada por el primer organismo de ensayo o parte del mismo y por el segundo organismo de ensayo o parte del mismo, para determinar de este modo la preferencia fenotípica del primer codón con respecto a la preferencia fenotípica del segundo codón en el organismo de interés o parte del mismo.

14. Un método de acuerdo con la reivindicación 13, que comprende adicionalmente determinar una comparación de preferencias fenotípicas de codones sinónimos individuales en el organismo de interés o parte del mismo.

15. Un método de construcción de un polinucleótido sintético a partir del cual se produce un polipéptido codificado a un nivel más alto en una célula de interés que a partir de un polinucleótido parental que codifica el mismo polipéptido, comprendiendo el método:

- determinar la eficiencia de la traducción de diferentes codones sinónimos en células del mismo tipo que la célula de interés, como se definen en la reivindicación 10, para determinar de este modo una comparación de eficiencias de traducción de codones sinónimos individuales en la célula de interés;

- seleccionar un primer codón del polinucleótido parental para el reemplazo por un codón sinónimo, en el que el codón sinónimo se selecciona basándose en que éste presenta una mayor eficiencia de traducción que el primer codón en la célula de interés de acuerdo con la comparación de las eficiencias de traducción; y

- reemplazar el primer codón por el codón sinónimo para construir el polinucleótido sintético.

16. Un método de acuerdo con la reivindicación 15, en el que el codón sinónimo se selecciona basándose en que éste corresponde a un codón de interrogación en una construcción sintética a partir de la cual se expresa la proteína indicadora en la célula de interés a un nivel que es al menos aproximadamente un 10 % superior al nivel de la proteína indicadora expresada a partir de una construcción sintética que comprende el primer codón como el codón de interrogación.

17. Un método de construcción de un polinucleótido sintético a partir del cual se produce un polipéptido codificado a un nivel más alto en una primera célula que en una segunda célula, comprendiendo el método:

- determinar la eficiencia de la traducción de diferentes codones sinónimos en células del mismo tipo que la primera célula y en células del mismo tipo que la segunda célula, como se define en la reivindicación 12, para determinar de este modo una comparación de eficiencias de traducción de codones sinónimos individuales entre la primera célula y la segunda célula;

- seleccionar un primer codón del polinucleótido parental para el reemplazo por un codón sinónimo, en el que el codón sinónimo se selecciona basándose en que éste presenta una mayor eficiencia de traducción en la primera célula que en la segunda célula de acuerdo con la comparación de las eficiencias de traducción; y

- reemplazar el primer codón por el codón sinónimo para construir el polinucleótido sintético.

18. Un método de acuerdo con la reivindicación 17, en el que el codón sinónimo es el mismo que el codón de interrogación en una construcción sintética a partir de la cual se expresa la proteína indicadora en la primera célula a un nivel que es al menos aproximadamente un 10 % superior al nivel de la proteína indicadora expresada a partir de la misma construcción sintética en la segunda célula.

19. Un método de construcción de un polinucleótido sintético a partir del cual se puede producir un polipéptido para conferir un fenotipo seleccionado a un organismo de interés o parte del mismo que difiere del conferido por un polinucleótido parental que codifica el mismo polipéptido, comprendiendo el método:

- determinar la preferencia de diferentes codones sinónimos para producir el fenotipo seleccionado en organismos de ensayo o partes de los mismos, como se define en la reivindicación 13, en el que los organismos de ensayo se seleccionan entre el grupo que consiste en un organismo de la misma especie que el organismo de interés y un organismo que está relacionado con el organismo de interés, para determinar de este modo una comparación de preferencias fenotípicas de codones sinónimos individuales en el organismo de interés;

- seleccionar un primer codón del polinucleótido parental para el reemplazo por un codón sinónimo, en el que el codón sinónimo se selecciona basándose en que éste presenta una preferencia fenotípica diferente a la del primer codón en la comparación de preferencias fenotípicas en el organismo o parte del mismo; y
- reemplazar el primer codón por el codón sinónimo para construir el polinucleótido sintético.

5 20. Un método de acuerdo con la reivindicación 19, en el que el polinucleótido sintético confiere el fenotipo seleccionado al organismo de interés o parte del mismo en una cualidad superior a la conferida por el polinucleótido parental.

10 21. Un método de acuerdo con la reivindicación 20, en el que el codón sinónimo se selecciona basándose en que éste corresponde a un codón de interrogación en una construcción sintética que confiere el fenotipo seleccionado en el organismo de interés o parte del mismo en una cualidad que es al menos aproximadamente un 10 % superior a la cualidad del fenotipo conferido por la construcción sintética que comprende el primer codón como el codón de

15 interrogación.

	1	11	21	31	41	51
IgkC1	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkS1-1	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkS1-2	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkS1-3	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkS1-4	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkC2	GGTACCGCCGCCACCATGGAGACCGACACCCCTCCTGCTGTTGGGTGCTGCTGCTCTGGGTT					
	61	71	81	91	101	111
IgkC1	CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG					
IgkS1-1	CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG					
IgkS1-2	CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG					
IgkS1-3	CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG					
IgkS1-4	CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG					
IgkC2	CCCGGCTCCACCGGCGACGGATCCATGCACGGCGACACCCACCCCTGCACGAGTACATG					
	121	131	141	151	161	171
IgkC1	TTAGATTGCAACCAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA					
IgkS1-1	TTAGATTGCAACCAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA					
IgkS1-2	TTAGATTGCAACCAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA					
IgkS1-3	TTAGATTGCAACCAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA					
IgkS1-4	TTAGATTGCAACCAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA					
IgkC2	CTGGACCTGCAGCCCGAGACCACCGGCTGTACGGCTACGGCCAGCTCAACGACGAGCAG					
	181	191	201	211	221	231
IgkC1	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCAATTAC					
IgkS1-1	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCAATTAC					
IgkS1-2	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCAATTAC					
IgkS1-3	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCTGAACCGGACAGAGCTCATTAC					
IgkS1-4	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCTGAACCGGACAGAGCCCAATTAC					
IgkC2	GAGGAGGAGGACGAGATCGACGGCCCGCCGGCCAGGCCGAGCCCGACCGCCCACTAC					
	241	251	261	271	281	291
IgkC1	AATATTGTAACCTTTTGTGCAAGTGTGACTCTACGCTTCGGTTGTGCGGTACAAAGCACA					
IgkS1-1	AATATTGTAACCTTTTGTGCAAGTGTGACTCTACGCTTCGGTTGTGCGGTACAAAGCACA					
IgkS1-2	AATATTGTAACCTTTTGTGCAAGTGTGACTCTACGCTTCGGTTGTGCGGTACAAAGCACA					
IgkS1-3	AATATTGTAACCTTTTGTGCAAGTGTGACTCTACGCTTCGGTTGTGCGGTACAAAGCACA					
IgkS1-4	AATATTGTAACCTTTTGTGCAAGTGTGACTCTACGCTTCGGTTGTGCGGTACAAAGCACA					
IgkC2	AACATCGTGACCTTCTGCTGCAAGTGTGACAGCACCCCTGCGCCTCTGCGGTGACAGCACC					
	301	311	321	331	341	351
IgkC1	CACGTAGACATTTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC					
IgkS1-1	CACGTAGACATTTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC					
IgkS1-2	CACGTAGACATTTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC					
IgkS1-3	CACGTAGACATTTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC					
IgkS1-4	CACGTAGACATTTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC					
IgkC2	CACGTGACATCCGACCCCTGGAGGACCTGCTGATGGGCACCCCTGGGCATCGTGTGCCCC					
	361	371	381			
IgkC1	ATCTGCTCTCAGAAGCCCTAAGAATTC					
IgkS1-1	ATCTGCTCTCAGAAGCCCTAAGAATTC					
IgkS1-2	ATCTGCTCTCAGAAGCCCTAAGAATTC					
IgkS1-3	ATCTGCTCTCAGAAGCCCTAAGAATTC					
IgkS1-4	ATCTGCTCTCAGAAGCCCTAAGAATTC					
IgkC2	ATCTGCTCCAGAAGCCCTAAGAATTC					

FIGURA 1

	1	11	21	31	41	51
IgkS1-5	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkS1-6	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkS1-7	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkS1-8	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkS1-9	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkS1-10	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkC1	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkC2	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
	61	71	81	91	101	111
IgkS1-5	CCAGGTTCCACTGGTGACGGATCCATGTCATGGAGATACACCTACATTGTCATGAATATATG					
IgkS1-6	CCAGGTTCCACTGGTGACGGATCCATGTCATGGAGATACACCTACATTGTCATGAATATATG					
IgkS1-7	CCAGGTTCCACTGGTGACGGATCCATGTCATGGAGATACACCTACATTGTCATGAATATATG					
IgkS1-8	CCAGGTTCCACTGGTGACGGATCCATGTCATGGAGATACACCTACATTGTCATGAATATATG					
IgkS1-9	CCAGGTTCCACTGGTGACGGATCCATGTCATGGAGATACACCTACATTGTCATGAATATATG					
IgkS1-10	CCAGGTTCCACTGGTGACGGATCCATGTCATGGAGATACACCTACATTGTCATGAATATATG					
IgkC1	CCAGGTTCCACTGGTGACGGATCCATGTCATGGAGATACACCTACATTGTCATGAATATATG					
IgkC2	CCAGGTTCCACTGGTGACGGATCCATGTCATGGAGATACACCTACATTGTCATGAATATATG					
	121	131	141	151	161	171
IgkS1-5	TTAGATTTCGAACCCAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA					
IgkS1-6	TTAGATTTCGAACCCAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA					
IgkS1-7	TTAGATTTCGAACCCAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA					
IgkS1-8	TTAGATTTCGAACCCAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA					
IgkS1-9	TTAGATTTCGAACCCAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA					
IgkS1-10	TTAGATTTCGAACCCAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA					
IgkC1	TTAGATTTCGAACCCAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA					
IgkC2	CTGGACCTGCAGCCGAGACCAACCGGCTGTACGGCTACGGCCAGCTCAACGACAGCAGC					
	181	191	201	211	221	231
IgkS1-5	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCCGGACAGGGCCCATAC					
IgkS1-6	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCCGGACAGAGCCCATAC					
IgkS1-7	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCCGGACAGAGCCCATAC					
IgkS1-8	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCCGGACAGAGCCCATAC					
IgkS1-9	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCCGGACAGAGCCCATAC					
IgkS1-10	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCCGGACAGAGCCCATAC					
IgkC1	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCCGGACAGAGCCCATAC					
IgkC2	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCCGGACAGAGCCCATAC					
	241	251	261	271	281	291
IgkS1-5	AATATTGTAACCTTTTGTGCAAGTGTGACTCTACGCTTAGGTTGTGCGTACAAAGCACA					
IgkS1-6	AATATTGTAACCTTTTGTGCAAGTGTGACTCTACGCTTAGGTTGTGCGTACAAAGCACA					
IgkS1-7	AATATTGTAACCTTTTGTGCAAGTGTGACTCTACGCTTAGGTTGTGCGTACAAAGCACA					
IgkS1-8	AATATTGTAACCTTTTGTGCAAGTGTGACTCTACGCTTAGGTTGTGCGTACAAAGCACA					
IgkS1-9	AATATTGTAACCTTTTGTGCAAGTGTGACTCTACGCTTAGGTTGTGCGTACAAAGCACA					
IgkS1-10	AATATTGTAACCTTTTGTGCAAGTGTGACTCTACGCTTAGGTTGTGCGTACAAAGCACA					
IgkC1	AATATTGTAACCTTTTGTGCAAGTGTGACTCTACGCTTAGGTTGTGCGTACAAAGCACA					
IgkC2	AACATCGTGACCTTCTGCTGCAAGTGTGACGACACCCCTGCGCTCTGCGTGCAGAGCACC					
	301	311	321	331	341	351
IgkS1-5	CACGTAGACATTAGGACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC					
IgkS1-6	CACGTAGACATTAGGACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC					
IgkS1-7	CACGTAGACATTAGGACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC					
IgkS1-8	CACGTAGACATTAGGACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC					
IgkS1-9	CACGTAGACATTAGGACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC					
IgkS1-10	CACGTAGACATTAGGACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC					
IgkC1	CACGTAGACATTAGGACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC					
IgkC2	CACGTAGACATTAGGACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC					
	361	371	381			
IgkS1-5	ATCTGCTCTCAGAAGCCCTAAGAATTC					
IgkS1-6	ATCTGCTCTCAGAAGCCCTAAGAATTC					
IgkS1-7	ATCTGCTCTCAGAAGCCCTAAGAATTC					
IgkS1-8	ATCTGCTCTCAGAAGCCCTAAGAATTC					
IgkS1-9	ATCTGCTCTCAGAAGCCCTAAGAATTC					
IgkS1-10	ATCTGCTCTCAGAAGCCCTAAGAATTC					
IgkC1	ATCTGCTCTCAGAAGCCCTAAGAATTC					
IgkC2	ATCTGCTCTCAGAAGCCCTAAGAATTC					

FIGURA 2

IgkC1	1	11	21	31	41	51
IgkS1-12	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkS1-31	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkC2	GGTACCGCCGCCACCATGGAGACAGACACCTCCTGCTGTGGGTGCTGCTGCTCTGGGTT					
IgkC1	61	71	81	91	101	111
IgkS1-12	CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG					
IgkS1-31	CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG					
IgkC2	CCCGGCTCCACCGCGACGGATCCATGCACGGCGACACCCACCCCTGCACAGTACATG					
IgkC1	121	131	141	151	161	171
IgkS1-12	TTAGATTGCAACCAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA					
IgkS1-31	TTAGATTGCAACCAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA					
IgkC2	CTGGACCTGCAGCCGAGACCACCGGCTGTACGGCTACGGCCAGCTCAACGACAGCAGC					
IgkC1	181	191	201	211	221	231
IgkS1-12	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC					
IgkS1-31	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC					
IgkC2	GAGGAGGAGGACGAGATCGACGGCCCCGCGGCCAGGCCGAGCCGACCGCGCCCACTAC					
IgkC1	241	251	261	271	281	291
IgkS1-12	AATATTGTAACCTTTTGTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA					
IgkS1-31	AATATTGTAACCTTTTGTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA					
IgkC2	AACATCGTGACCTTCTGCTGCAAGTGCAGACAGCACCTGCGCCTCTGCGTGCAGAGCACC					
IgkC1	301	311	321	331	341	351
IgkS1-12	CACGTAGACATTGCTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC					
IgkS1-31	CACGTAGACATTGCTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC					
IgkC2	CACGTGGACATCCGCACCCCTGGAGGACCTGCTGATGGGCACCTGGGACATCGTGTGCCCC					
IgkC1	361	371	381			
IgkS1-12	ATCTGCTCTCAGAAGCCCTAAGAATTTC					
IgkS1-31	ATCTGCTCTCAGAAGCCCTAAGAATTTC					
IgkC2	ATCTGCTCCAGAAGCCCTAAGAATTTC					

FIGURA 3

	1	11	21	31	41	51
IgkC1	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkS1-13	GGTACCGCCGCCACCATGGAGACAGATACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkS1-14	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkC2	GGTACCGCCGCCACCATGGAGACCGACACCTCCTGCTGTGGGTGCTGCTGCTCTGGGTT					
	61	71	81	91	101	111
IgkC1	CCAGGTTCCACTGGTGACGGATCCATGTCATGGAGATACACCTACATTGCATGAATATATG					
IgkS1-13	CCAGGTTCCACTGGTGATGGATCCATGTCATGGAGATACACCTACATTGCATGAATATATG					
IgkS1-14	CCAGGTTCCACTGGTGACGGATCCATGTCATGGAGACACCTACATTGCATGAATATATG					
IgkC2	CCCGGCTCCACCGCGACGGATCCATGTCACGGCGACACCCCCACCCTGCACGAGTACATG					
	121	131	141	151	161	171
IgkC1	TTAGATTTCGAACCAAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA					
IgkS1-13	TTAGATTTCGAACCAAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGATAGCTCA					
IgkS1-14	TTAGACTTCGAACCAAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA					
IgkC2	CTGGACCTGCAGCCCGAGACCAACCGGCTGTACGGCTACGGCCAGCTCAACGACAGCAGC					
	181	191	201	211	221	231
IgkC1	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCCATTAC					
IgkS1-13	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGATAGAGCCCCATTAC					
IgkS1-14	GAGGAGGAGGACGAAATAGACGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCCATTAC					
IgkC2	GAGGAGGAGGACGAGATCGACGGCCCCCGCGCCAGGCCGAGCCCCACCGCGCCCCACTAC					
	241	251	261	271	281	291
IgkC1	AATATTGTAACCTTTTGTGCAAGTGTGACTCTACGCTTCGGTTGTGCGGTACAAAGCACA					
IgkS1-13	AATATTGTAACCTTTTGTGCAAGTGTGATTCTACGCTTCGGTTGTGCGGTACAAAGCACA					
IgkS1-14	AATATTGTAACCTTTTGTGCAAGTGTGACTCTACGCTTCGGTTGTGCGGTACAAAGCACA					
IgkC2	AACATCGTGACCTTCTGCTGCAAGTGTGCGACAGCACCTGCGCCTCTGCGTGCAGAGCACC					
	301	311	321	331	341	351
IgkC1	CACGTAGACATTTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC					
IgkS1-13	CACGTAGATATTTCGTACTTTGGAAGATCTGTTAATGGGCACACTAGGAATTGTGTGCCCC					
IgkS1-14	CACGTAGACATTTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC					
IgkC2	CACGTGGACATCCGACCCCTGGAGGACCTGCTGATGGGCACCCCTGGGCATCGTGTGCCCC					
	361	371	381			
IgkC1	ATCTGCTCTCAGAAGCCCTAAGAATTC					
IgkS1-13	ATCTGCTCTCAGAAGCCCTAAGAATTC					
IgkS1-14	ATCTGCTCTCAGAAGCCCTAAGAATTC					
IgkC2	ATCTGCTCCAGAAGCCCTAAGAATTC					

FIGURA 4

IgkC1	1	11	21	31	41	51
IgkS1-15	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkS1-16	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkC2	GGTACCGCCGCCACCATGGAGACCGACACCCCTCCTGCTGTGGGTGCTGCTGCTCTGGGTT					
IgkC1	61	71	81	91	101	111
IgkS1-15	CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG					
IgkS1-16	CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG					
IgkC2	CCCGGCTCCACCGCGACGGATCCATGCACGGCGACACCCCCACCTGCACGAGTACATG					
IgkC1	121	131	141	151	161	171
IgkS1-15	TTAGATTGCAACCAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA					
IgkS1-16	TTAGATTGCAACCAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA					
IgkC2	CTGGACCTGCAGCCCGAGACCACCGGCTGTACGGCTACGGCCAGCTCAACGACAGCAGC					
IgkC1	181	191	201	211	221	231
IgkS1-15	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC					
IgkS1-16	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC					
IgkC2	GAGGAGGAGGACGAGATCGACGGCCCCGCGGCCAGGCCGAGCCGACCGGCCCACTAC					
IgkC1	241	251	261	271	281	291
IgkS1-15	AATATTGTAACCTTTTGTGCAAGTGTGACTCTACGCITTCGGTTGTGCGTACAAAGCACA					
IgkS1-16	AATATTGTAACCTTTTGTGCAAGTGTGACTCTACGCITTCGGTTGTGCGTACAAAGCACA					
IgkC2	AACATCGTGACCTTCTGCTGCAAGTGTGACAGCACCCTGCGCCTCTGCGTGCAGAGCACC					
IgkC1	301	311	321	331	341	351
IgkS1-15	CACGTAGACATTCTGACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC					
IgkS1-16	CACGTAGACATTCTGACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGTCCC					
IgkC2	CACGTAGACATTCTGACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGTCCC					
IgkC1	361	371	381			
IgkS1-15	ATCTGCTCTCAGAAGCCCTAAGAATTC					
IgkS1-16	ATCTGTTCTCAGAAGCCCTAAGAATTC					
IgkC2	ATCTGCTCTCAGAAGCCCTAAGAATTC					

FIGURA 5

	1	11	21	31	41	51
IgkS1-17	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkS1-18	GGTACCGCCGCCACCATGGAAACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkC2	GGTACCGCCGCCACCATGGAGACCGACACCCCTCCTGCTGCTGGGTGCTGCTGCTCTGGGTG					
IgkC1	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
	61	71	81	91	101	111
IgkS1-17	CCAGGTTCCACTGGTGACGGATCCATGTCATGGAGATACACCTACATTGTCATGAGTATATG					
IgkS1-18	CCAGGTTCCACTGGTGACGGATCCATGTCATGGAGATACACCTACATTGTCATGAATATATG					
IgkC2	CCCGGCTCCACCGGCGACGGATCCATGTCACGGCGACACCCCACTGTCACGAGTACATG					
IgkC1	CCAGGTTCCACTGGTGACGGATCCATGTCATGGAGATACACCTACATTGTCATGAATATATG					
	121	131	141	151	161	171
IgkS1-17	TTAGATTTGCAACCAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA					
IgkS1-18	TTAGATTTGCAACCAGAAACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA					
IgkC2	CTGGACCTGTCAGCCCGAGACCACCGGCCTGTACGGCTACGGCCAGCTCAACGACAGCAGC					
IgkC1	TTAGATTTGCAACCAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA					
	181	191	201	211	221	231
IgkS1-17	GAGGAGGAGGATGAGATAGATGGTCCAGCTGGACAAGCAGAGCCGGACAGAGCCCATTAC					
IgkS1-18	GAAGAAGAAGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC					
IgkC2	GAGGAGGAGGACGAGATCGACGGCCCGCCGGCCAGGCCGAGCCGACCGGCCCACTAC					
IgkC1	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC					
	241	251	261	271	281	291
IgkS1-17	AATATGTAAACCTTTTGTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA					
IgkS1-18	AATATGTAAACCTTTTGTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA					
IgkC2	AACATCGTGACCTTCTGCTGCAAGTGCAGACAGCACCTGCGCCTCTGCGTGCAGAGCACC					
IgkC1	AATATGTAAACCTTTTGTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA					
	301	311	321	331	341	351
IgkS1-17	CACGTAGACATTGCTACTTTGGAGGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC					
IgkS1-18	CACGTAGACATTGCTACTTTGGAGGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC					
IgkC2	CACGTGACATCCGCAACCCTGGAGGACCTGCTGATGGGCACCTGGGCATCGTGTGCCCC					
IgkC1	CACGTAGACATTGCTACTTTGGAGGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC					
	361	371	381			
IgkS1-17	ATCTGCTCTCAGAAGCCCTAAGAATTC					
IgkS1-18	ATCTGCTCTCAGAAGCCCTAAGAATTC					
IgkC2	ATCTGCTCCCAGAAGCCCTAAGAATTC					
IgkC1	ATCTGCTCTCAGAAGCCCTAAGAATTC					

FIGURA 6

IgkC1	1	11	21	31	41	51
IgkS1-19	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkS1-20	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkC2	GGTACCGCCGCCACCATGGAGACCGACACCCCTCCTGCTGTGGGTGCTGCTGCTCTGGGTT					
IgkC1	61	71	81	91	101	111
IgkS1-19	CCAGGTTCCACTGGTGACGGATCCATGTCATGGAGATACACCTACATTGCATGAATATATG					
IgkS1-20	CCAGGTTCCACTGGTGACGGATCCATGTCATGGAGATACACCTACATTGCATGAATATATG					
IgkC2	CCCGGCTCCACCGGCGACGGATCCATGCACGGCGACACCCCCACCTGCACGAGTACATG					
IgkC1	121	131	141	151	161	171
IgkS1-19	TTAGATTTCGAACCCAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA					
IgkS1-20	TTAGATTTCGAACCCAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA					
IgkC2	CTGGACCTGCAGCCCGAGACCAACGGCTGTACGGCTACGGCCAGCTCAACGACAGCAGC					
IgkC1	181	191	201	211	221	231
IgkS1-19	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC					
IgkS1-20	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC					
IgkC2	GAGGAGGAGGACGAGATCGACGGCCCCGCCGCCAGGCCGAGCCGACCGCGCCCACTAC					
IgkC1	241	251	261	271	281	291
IgkS1-19	AATATTGTAACCTTTTGTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA					
IgkS1-20	AATATTGTAACCTTTTGTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA					
IgkC2	AACATCGTGACCTTTCTGTGCAAGTGTGACAGCACCTTGCCTCTGCGTGCAGAGCACC					
IgkC1	301	311	321	331	341	351
IgkS1-19	CACGTAGACATTCGTACTTTGGAAGACCTGTAAATGGGCACACTAGGAATTGTGTGCCCC					
IgkS1-20	CACGTAGACATTCGTACTTTGGAAGACCTGTAAATGGGCACACTAGGAATTGTGTGCCCC					
IgkC2	CACGTGGACATCCGCACCTTGGAGGACCTGTGTATGGGCACCTGGGCATCGTGTGCCCC					
IgkC1	361	371	381			
IgkS1-19	ATCTGCTCTCAGAAGCCCTAAGAATTC					
IgkS1-20	ATCTGCTCTCAGAAGCCCTAAGAATTC					
IgkC2	ATCTGCTCCAGAAGCCCTAAGAATTC					

FIGURA 7

IgkC1	1	11	21	31	41	51
IgkS1-21	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkS1-22	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkS1-23	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkS1-24	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkC2	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkC1	61	71	81	91	101	111
IgkS1-21	CCAGGTTCCACTGGTGACGGATCCATGTCATGGAGATACACCTACATTGTCATGAATATATG					
IgkS1-22	CCAGGTTCCACTGGTGACGGATCCATGTCATGGAGATACACCTACATTGTCATGAATATATG					
IgkS1-23	CCAGGTTCCACTGGTGACGGATCCATGTCATGGAGATACACCTACATTGTCATGAATATATG					
IgkS1-24	CCAGGTTCCACTGGTGACGGATCCATGTCATGGAGATACACCTACATTGTCATGAATATATG					
IgkC2	CCAGGTTCCACTGGTGACGGATCCATGTCATGGAGATACACCTACATTGTCATGAATATATG					
IgkC1	121	131	141	151	161	171
IgkS1-21	TTAGATTGTCAACCAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA					
IgkS1-22	TTAGATTGTCAACCAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA					
IgkS1-23	TTAGATTGTCAACCAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA					
IgkS1-24	TTAGATTGTCAACCAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA					
IgkC2	TTAGATTGTCAACCAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA					
IgkC1	181	191	201	211	221	231
IgkS1-21	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCCATTAC					
IgkS1-22	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCCATTAC					
IgkS1-23	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCCATTAC					
IgkS1-24	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCCATTAC					
IgkC2	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCCATTAC					
IgkC1	241	251	261	271	281	291
IgkS1-21	AATATTGTAAACCTTTTGTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA					
IgkS1-22	AATATTGTAAACCTTTTGTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA					
IgkS1-23	AATATTGTAAACCTTTTGTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA					
IgkS1-24	AATATTGTAAACCTTTTGTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA					
IgkC2	AATATTGTAAACCTTTTGTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA					
IgkC1	301	311	321	331	341	351
IgkS1-21	CACGTAGACATTTCGTACTTTGGAAGACCTGTTAATGGGACACTAGGGAATTGTGTGCCCC					
IgkS1-22	CACGTAGACATTTCGTACTTTGGAAGACCTGTTAATGGGACACTAGGGAATTGTGTGCCCC					
IgkS1-23	CACGTAGACATTTCGTACTTTGGAAGACCTGTTAATGGGACACTAGGGAATTGTGTGCCCC					
IgkS1-24	CACGTAGACATTTCGTACTTTGGAAGACCTGTTAATGGGACACTAGGGAATTGTGTGCCCC					
IgkC2	CACGTAGACATTTCGTACTTTGGAAGACCTGTTAATGGGACACTAGGGAATTGTGTGCCCC					
IgkC1	361	371	381			
IgkS1-21	ATCTGCTCTCAGAAGCCCTAAGAATTC					
IgkS1-22	ATCTGCTCTCAGAAGCCCTAAGAATTC					
IgkS1-23	ATCTGCTCTCAGAAGCCCTAAGAATTC					
IgkS1-24	ATCTGCTCTCAGAAGCCCTAAGAATTC					
IgkC2	ATCTGCTCTCAGAAGCCCTAAGAATTC					

FIGURA 8

IgkC1	1	11	21	31	41	51
IgkS1-25	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkS1-26	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkC2	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkC1	61	71	81	91	101	111
IgkS1-25	CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG					
IgkS1-26	CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG					
IgkC2	CCCGGCTCCACCGCGACGGATCCATGCACGGCGACACCCACCCCTGCACGAGTACATG					
IgkC1	121	131	141	151	161	171
IgkS1-25	TTAGATTTGCAACCAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA					
IgkS1-26	TTAGATTTGCAACCAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA					
IgkC2	CTGGACCTGCAGCCCGAGACCACCGGCTGTACGGCTACGGCCAGCTCAACGACAGCAGC					
IgkC1	181	191	201	211	221	231
IgkS1-25	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC					
IgkS1-26	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC					
IgkC2	GAGGAGGAGGACGAGATCGACGGCCCCGCCGCCAGGCCGAGCCCGACCGCGCCCACTAC					
IgkC1	241	251	261	271	281	291
IgkS1-25	AATATTGTAACCTTTTGTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA					
IgkS1-26	AATATTGTAACCTTTTGTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA					
IgkC2	AACATCGTGACCTTCTGCTGCAAGTGTGCGACAGCACCTGCGCCTCTGCGTGCAAGCACC					
IgkC1	301	311	321	331	341	351
IgkS1-25	CACGTAGACATTTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC					
IgkS1-26	CACGTAGACATTTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC					
IgkC2	CACGTGGACATCCGCACCCCTGGAGGACCTGCTGATGGGCACCCCTGGGCATCGTGTGCCCC					
IgkC1	361	371	381			
IgkS1-25	ATCTGCTCTCAGAAGCCCTAAGAATTC					
IgkS1-26	ATCTGCTCTCAGAAGCCCTAAGAATTC					
IgkC2	ATCTGCTCCAGAAGCCCTAAGAATTC					

FIGURA 9

	1	11	21	31	41	51
IgkC1	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkS1-27	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkS1-28	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkS1-29	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkC2	GGTACCGCCGCCACCATGGAGACCGACACCTCCTGCTGTTGGGTGCTGCTGCTCTGGGTT					
	61	71	81	91	101	111
IgkC1	CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG					
IgkS1-27	CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG					
IgkS1-28	CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG					
IgkS1-29	CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG					
IgkC2	CCCGGCTCCACCGCGACGGATCCATGCACGGCGACACCCACCTGCACGAGTACATG					
	121	131	141	151	161	171
IgkC1	TTAGATTGCAACCAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA					
IgkS1-27	TTAGATTGCAACCAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA					
IgkS1-28	TTAGATTGCAACCAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA					
IgkS1-29	TTAGATTGCAACCAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA					
IgkC2	CTGGACCTGCAGCCGAGACCACCGGCTGTACGGCTACGGCCAGCTCAACGACAGCAGC					
	181	191	201	211	221	231
IgkC1	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC					
IgkS1-27	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC					
IgkS1-28	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC					
IgkS1-29	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC					
IgkC2	GAGGAGGAGGACGAGATCGACGGCCCGCCGGCCAGGCCGAGCCGACCGCGCCCACTAC					
	241	251	261	271	281	291
IgkC1	AATATTGTAACCTTTTGTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA					
IgkS1-27	AATATTGTAACCTTTTGTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA					
IgkS1-28	AATATTGTAACCTTTTGTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA					
IgkS1-29	AATATTGTAACCTTTTGTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA					
IgkC2	AACATCGTGACCTTCTGCTGCAAGTGGACAGCACCTGCGCCTCTGCGTGAGAGCACC					
	301	311	321	331	341	351
IgkC1	CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC					
IgkS1-27	CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC					
IgkS1-28	CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC					
IgkS1-29	CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC					
IgkC2	CACGTGGACATCCGACCCCTGGAGGACCTGCTGATGGGCACCTGGGCATCGTGTGCCCC					
	361	371	381			
IgkC1	ATCTGCTCTCAGAAGCCCTAAGAATTC					
IgkS1-27	ATATGCTCTCAGAAGCCCTAAGAATTC					
IgkS1-28	ATTTGCTCTCAGAAGCCCTAAGAATTC					
IgkS1-29	ATCTGCTCTCAGAAGCCCTAAGAATTC					
IgkC2	ATCTGCTCCAGAAGCCCTAAGAATTC					

FIGURA 10

	1	11	21	31	41	51
IgkS1-50	GGTACCGCCGCCACCATGGAACTGACACTCTGCTGCTGTTGGGTACTGCTGCTGTTGGGTT					
IgkS1-51	GGTACCGCCGCCACCATGGAACTGACACTCTACTACTATGGGTACTACTACTATGGGTT					
IgkS1-52	GGTACCGCCGCCACCATGGAACTGACACTCTTCTTCTTGGGTACTTCTTCTTGGGTT					
IgkS1-53	GGTACCGCCGCCACCATGGAACTGACACTCTCTCTCTGTTGGGTACTCTCTCTGTTGGGTT					
IgkS1-54	GGTACCGCCGCCACCATGGAACTGACACTTTGTGTGTGTGGGTATTGTGTGTGTGGGTT					
IgkS1-55	GGTACCGCCGCCACCATGGAACTGACACTTTATTATTATGGGTATTATTATTATGGGTT					
IgkC3	GGTACCGCCGCCACCATGGAGACGACACACTCTCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkC4	GGTACCGCCGCCACCATGGAGACCGACACCTCTCTGCTGTGGGTGCTGCTGCTCTGGGTT					
	61	71	81	91	101	111
IgkS1-50	CCAGGATCGACTGGAGACGGATCCATGTCATGGAGACACTCCAACCTGTCATGAATATATG					
IgkS1-51	CCAGGATCGACTGGAGACGGATCCATGTCATGGAGACACTCCAACCTACATGAATATATG					
IgkS1-52	CCAGGATCGACTGGAGACGGATCCATGTCATGGAGACACTCCAACCTTTCATGAATATATG					
IgkS1-53	CCAGGATCGACTGGAGACGGATCCATGTCATGGAGACACTCCAACCTCCATGAATATATG					
IgkS1-54	CCAGGATCGACTGGAGACGGATCCATGTCATGGAGACACTCCAACCTTTCATGAATATATG					
IgkS1-55	CCAGGATCGACTGGAGACGGATCCATGTCATGGAGACACTCCAACCTTTCATGAATATATG					
IgkC3	CCAGGTTCCACTGGTGAACGGATCCATGTCATGGAGATACCTACATTGTCATGAATATATG					
IgkC4	CCCGGCTCCACCGGCGACGGATCCATGTCACGGCGACACCCCACTCTGACGAGTACATG					
	121	131	141	151	161	171
IgkS1-50	CTGGACCTGCAACCGGAACTACTGACCTGTACTGCTATGAACAACTGAATGACAGCTCG					
IgkS1-51	CTAGACCTTACAACCGGAACTACTGACCTATATGCTATGAACAACTAAATGACAGCTCG					
IgkS1-52	CTTGACCTTCAACCGGAACTACTGACCTTTACTGCTATGAACAACTTAATGACAGCTCG					
IgkS1-53	CTCGACCTTCAACCGGAACTACTGACCTCTACTGCTATGAACAACTCAATGACAGCTCG					
IgkS1-54	TTGGACTTGAACCGGAACTACTGACTTGTACTGCTATGAACAACTGAATGACAGCTCG					
IgkS1-55	TTAGACTTACAACCGGAACTACTGACTTATCTGCTATGAACAACTTAATGACAGCTCG					
IgkC3	TTAGATTTCGAACCGAGACAACTGATCTCTACTGTTATGAGCAATTAATGACAGCTCA					
IgkC4	CTGGACCTGCAACCGGAGACCAACCGACCTGTACTGCTACGAGCAGCTCAACGACAGCAGC					
	181	191	201	211	221	231
IgkS1-50	GAAGAAGAAGACGAAATAGACGGACCTGCAGGACAAGCAGAACCGAGCCGCGCATTAC					
IgkS1-51	GAAGAAGAAGACGAAATAGACGGACCTGCAGGACAAGCAGAACCGAGCCGCGCATTAC					
IgkS1-52	GAAGAAGAAGACGAAATAGACGGACCTGCAGGACAAGCAGAACCGAGCCGCGCATTAC					
IgkS1-53	GAAGAAGAAGACGAAATAGACGGACCTGCAGGACAAGCAGAACCGAGCCGCGCATTAC					
IgkS1-54	GAAGAAGAAGACGAAATAGACGGACCTGCAGGACAAGCAGAACCGAGCCGCGCATTAC					
IgkS1-55	GAAGAAGAAGACGAAATAGACGGACCTGCAGGACAAGCAGAACCGAGCCGCGCATTAC					
IgkC3	GAGGAGGAGGATGAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCACTAC					
IgkC4	GAGGAGGAGGACGAGATCGAGGCCCCCGCCGCGGAGCCGAGCCGAGCCGAGCCCACTAC					
	241	251	261	271	281	291
IgkS1-50	AATATTGTAACCTTTTGTCTGCAAGTGCAGCAGTACTCTGCGCCTGTGCGTACAAAGCACT					
IgkS1-51	AATATTGTAACCTTTTGTCTGCAAGTGCAGCAGTACTCTACGCCCTATGCGTACAAAGCACT					
IgkS1-52	AATATTGTAACCTTTTGTCTGCAAGTGCAGCAGTACTCTTTCGCTTTGCGTACAAAGCACT					
IgkS1-53	AATATTGTAACCTTTTGTCTGCAAGTGCAGCAGTACTCTCGCCCTCTGCGTACAAAGCACT					
IgkS1-54	AATATTGTAACCTTTTGTCTGCAAGTGCAGCAGTACTTTGCGCTTGTGCGTACAAAGCACT					
IgkS1-55	AATATTGTAACCTTTTGTCTGCAAGTGCAGCAGTACTTACGCTTATGCGTACAAAGCACT					
IgkC3	AATATTGTAACCTTTTGTCTGCAAGTGCAGCAGTACTACGCTTTCGCTTGTGCGTACAAAGCA					
IgkC4	AACATCGTGACCTTCTGCTGCAAGTGCAGCAGCACCTCTGCGCCTCTGCGTGCAGAGCACC					
	301	311	321	331	341	351
IgkS1-50	CATGTAGACATTGCGCACTCTGGAAGACCTGCTGATGGGAACCTTGGGAATTGTTTGCCCG					
IgkS1-51	CATGTAGACATTGCGCACTCTAGAAGACCTACTAATGGGAACCTTAGGAATTGTTTGCCCG					
IgkS1-52	CATGTAGACATTGCGCACTCTTGAAGACCTTCTTATGGGAACCTTGGGAATTGTTTGCCCG					
IgkS1-53	CATGTAGACATTGCGCACTCTGGAAGACCTCTCATGGGAACCTTGGGAATTGTTTGCCCG					
IgkS1-54	CATGTAGACATTGCGCACTTTTGAAGACTTGTGATGGGAACCTTGGGAATTGTTTGCCCG					
IgkS1-55	CATGTAGACATTGCGCACTTTAGAAGACTTATTAATGGGAACCTTAGGAATTGTTTGCCCG					
IgkC3	CACGTAGACATTGCTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTTTGCCCG					
IgkC4	CACGTGACATCCGACCCCTGGAGGACCTGCTGATGGGCACCTTGGGCATCGTGTGCCCC					
	361	371	381			
IgkS1-51	ATCTGCTCGCAAAAGCCTTAAGAATTC					
IgkS1-53	ATCTGCTCGCAAAAGCCTTAAGAATTC					
IgkS1-50	ATCTGCTCGCAAAAGCCTTAAGAATTC					
IgkS1-52	ATCTGCTCGCAAAAGCCTTAAGAATTC					
IgkS1-55	ATCTGCTCGCAAAAGCCTTAAGAATTC					
IgkS1-54	ATCTGCTCGCAAAAGCCTTAAGAATTC					
IgkC3	ATCTGCTCTCAGAAGCCTTAAGAATTC					
IgkC4	ATCTGCTCCCAAGCCTTAAGAATTC					

FIGURA 11

	1	11	21	31	41	51
IgkS1-32	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkS1-33	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkC1	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkC2	GGTACCGCCGCCACCATGGAGACAGACACCTCCTGCTGCTGCTGCTGCTCTGGGTT					
	61	71	81	91	101	111
IgkS1-32	CCAGGTTCCACTGGTGACGGATCCATGTCATGGAGATACACCTACATTGCATGAATATATG					
IgkS1-33	CCAGGTTCCACTGGTGACGGATCCATGTCATGGAGATACACCTACATTGCATGAATATATG					
IgkC1	CCAGGTTCCACTGGTGACGGATCCATGTCATGGAGATACACCTACATTGCATGAATATATG					
IgkC2	CCCGGCTCCACCGGCGACGGATCCATGCACGGCGACACCCCACTGCACGAGTACATG					
	121	131	141	151	161	171
IgkS1-32	TTAGATTTTCAACCGAGACAACCTGGTTTTTACGGTTATGGGCAATTAAATGACAGCTCA					
IgkS1-33	TTAGATTTTCAACCGAGACAACCTGGTTTTTACGGTTATGGGCAATTAAATGACAGCTCA					
IgkC1	TTAGATTTGCAACCGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA					
IgkC2	CTGGACCTGCAGCCGAGACCAACCGGCTGTACGGCTACGGCCAGCTCAACGACAGCAGC					
	181	191	201	211	221	231
IgkS1-32	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATAC					
IgkS1-33	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATAC					
IgkC1	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATAC					
IgkC2	GAGGAGGAGGACGAGATCGACGGCCCGCCGGCCAGCCGAGCCCGACCGCCGCCCCACTAC					
	241	251	261	271	281	291
IgkS1-32	AATATTGTAAACCTTTTGTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA					
IgkS1-33	AATATTGTAAACCTTTTGTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA					
IgkC1	AATATTGTAAACCTTTTGTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA					
IgkC2	AACATCGTGACCTTCTGCTGCAAGTGCAGACAGCACCTGCGCTCTGCGTGACAGACACC					
	301	311	321	331	341	351
IgkS1-32	CACGTAGACATTGCTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC					
IgkS1-33	CACGTAGACATTGCTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC					
IgkC1	CACGTAGACATTGCTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC					
IgkC2	CACGTGGACATCCGACCCCTGGAGACCTGCTGATGGGCACCTGGGCATCGTGTGCCCC					
	361	371	381			
IgkS1-32	ATCTGCTCTCAGAAGCCCTAAGAATTC					
IgkS1-33	ATCTGCTCTCAGAAGCCCTAAGAATTC					
IgkC1	ATCTGCTCTCAGAAGCCCTAAGAATTC					
IgkC2	ATCTGCTCCAGAAGCCCTAAGAATTC					

FIGURA 12

	1	11	21	31	41	51
IgkS1-56	GGTACCGCCGCCACCATGGAAACTGACACTCTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkS1-57	GGTACCGCCGCCACCATGGAAACTGACACTCTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkS1-58	GGTACCGCCGCCACCATGGAAACTGACACTCTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkS1-59	GGTACCGCCGCCACCATGGAAACTGACACTCTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkC3	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkC4	GGTACCGCCGCCACCATGGAGACCGACCCCTCCTGCTGTGGGTGCTGCTGCTCTGGGTT					
	61	71	81	91	101	111
IgkS1-56	CCGGGATCGACTGGAGACGGATCCATGTCATGGAGACACTCCGACTTTGCATGAATATATG					
IgkS1-57	CCAGGATCGACTGGAGACGGATCCATGTCATGGAGACACTCCAACTTTGCATGAATATATG					
IgkS1-58	CCTGGATCGACTGGAGACGGATCCATGTCATGGAGACACTCCCTACTTTGCATGAATATATG					
IgkS1-59	CCCGGATCGACTGGAGACGGATCCATGTCATGGAGACACTCCCACTTTGCATGAATATATG					
IgkC3	CCAGGTTCCACTGGTGACGGATCCATGTCATGGAGATACACCTACATTGCATGAATATATG					
IgkC4	CCCGGCTCCACCGCGACGGATCCATGCACGGCGACACCCCACTTGCACGAGTACATG					
	121	131	141	151	161	171
IgkS1-56	CTCGACTTGCAACCGGAACTACTGACCTCTACTGCTATGAACAATTGAATGACAGCTCG					
IgkS1-57	CTCGACTTGCAACCGGAACTACTGACCTCTACTGCTATGAACAATTGAATGACAGCTCG					
IgkS1-58	CTCGACTTGCAACCTGAACTACTGACCTCTACTGCTATGAACAATTGAATGACAGCTCG					
IgkS1-59	CTCGACTTGCAACCGGAACTACTGACCTCTACTGCTATGAACAATTGAATGACAGCTCG					
IgkC3	TTAGATTGCAACCGAGAGCAACTGATCTCTACTGTTATGAGCAATTAAATGACAGCTCA					
IgkC4	CTGGACCTGCAGCCCGAGACCACCGACCTGTACTGTACGAGCAGCTCAACGACAGCAGC					
	181	191	201	211	221	231
IgkS1-56	GAAGAAGAAGACGAAATAGACGGACCGGACGAGACAGCAACCGGACCGCGCACATTAC					
IgkS1-57	GAAGAAGAAGACGAAATAGACGGACCGGACGAGACAGCAACCGGACCGCGCACATTAC					
IgkS1-58	GAAGAAGAAGACGAAATAGACGGACCGGACGAGACAGCAACCTGACCGCGCACATTAC					
IgkS1-59	GAAGAAGAAGACGAAATAGACGGACCGGACGAGACAGCAACCGGACCGCGCACATTAC					
IgkC3	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAGCAGAACCGGACAGACCCATTAC					
IgkC4	GAGGAGGAGGACGAGATCGACGGCCCCGCGCCAGGCCGAGCCGACCGCGCCCACTAC					
	241	251	261	271	281	291
IgkS1-56	AATATTGTAACCTTTTGTGCAAGTGCGACAGTACTCTCCGCTTGTGCGTACAAAGCACT					
IgkS1-57	AATATTGTAACCTTTTGTGCAAGTGCGACAGTACTCTCCGCTTGTGCGTACAAAGCACT					
IgkS1-58	AATATTGTAACCTTTTGTGCAAGTGCGACAGTACTCTCCGCTTGTGCGTACAAAGCACT					
IgkS1-59	AATATTGTAACCTTTTGTGCAAGTGCGACAGTACTCTCCGCTTGTGCGTACAAAGCACT					
IgkC3	AATATTGTAACCTTTTGTGCAAGTGCGACTCTACGCTTCCGCTTGTGCGTACAAAGCACA					
IgkC4	AACATCGTGACCTTCTGCTGCAAGTGCGACAGCACCTCGCGCTCTGCGTGACAGCACC					
	301	311	321	331	341	351
IgkS1-56	CATGTAGACATTTCGCACCTTTGGAAGACCTCCTCATGGGAACCTTTGGGAATTGTTTGC					
IgkS1-57	CATGTAGACATTTCGCACCTTTGGAAGACCTCCTCATGGGAACCTTTGGGAATTGTTTGC					
IgkS1-58	CATGTAGACATTTCGCACCTTTGGAAGACCTCCTCATGGGAACCTTTGGGAATTGTTTGC					
IgkS1-59	CATGTAGACATTTCGCACCTTTGGAAGACCTCCTCATGGGAACCTTTGGGAATTGTTTGC					
IgkC3	CACGTAGACATTTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTTTGC					
IgkC4	CACGTGGACATCCGCACCTTGGAGGACCTGCTGATGGGCACCTTGGGCATCGTGTGCCCC					
	361	371	381			
IgkS1-56	ATCTGCTCGCAAAAGCCGTAAGAATTC					
IgkS1-57	ATCTGCTCGCAAAAGCCGTAAGAATTC					
IgkS1-58	ATCTGCTCGCAAAAGCCGTAAGAATTC					
IgkS1-59	ATCTGCTCGCAAAAGCCGTAAGAATTC					
IgkC3	ATCTGCTCTCAGAAGCCCTAAGAATTC					
IgkC4	ATCTGCTCCAGAAGCCCTAAGAATTC					

FIGURA 13

IgkS1-34	1	11	21	31	41	51
IgkS1-35	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkS1-36	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkS1-37	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkS1-38	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkS1-39	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkC1	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkC2	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkS1-34	61	71	81	91	101	111
IgkS1-35	CCAGGTAGTACTGGTGACGGAAGTATGCATGGAGATACACCTACATTGCATGAATATATG					
IgkS1-36	CCAGGTAGTACTGGTGACGGAAGTATGCATGGAGATACACCTACATTGCATGAATATATG					
IgkS1-37	CCAGGTAGTACTGGTGACGGAAGTATGCATGGAGATACACCTACATTGCATGAATATATG					
IgkS1-38	CCAGGTAGTACTGGTGACGGAAGTATGCATGGAGATACACCTACATTGCATGAATATATG					
IgkS1-39	CCAGGTAGTACTGGTGACGGAAGTATGCATGGAGATACACCTACATTGCATGAATATATG					
IgkC1	CCAGGTAGTACTGGTGACGGAAGTATGCATGGAGATACACCTACATTGCATGAATATATG					
IgkC2	CCAGGTAGTACTGGTGACGGAAGTATGCATGGAGATACACCTACATTGCATGAATATATG					
IgkS1-34	121	131	141	151	161	171
IgkS1-35	TTAGATTGCAACCCAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGTAGT					
IgkS1-36	TTAGATTGCAACCCAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGTAGT					
IgkS1-37	TTAGATTGCAACCCAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGTAGT					
IgkS1-38	TTAGATTGCAACCCAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGTAGT					
IgkS1-39	TTAGATTGCAACCCAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGTAGT					
IgkC1	TTAGATTGCAACCCAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGTAGT					
IgkC2	TTAGATTGCAACCCAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGTAGT					
IgkS1-34	181	191	201	211	221	231
IgkS1-35	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAAGCAGAACCGGACAGAGCCCATAC					
IgkS1-36	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAAGCAGAACCGGACAGAGCCCATAC					
IgkS1-37	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAAGCAGAACCGGACAGAGCCCATAC					
IgkS1-38	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAAGCAGAACCGGACAGAGCCCATAC					
IgkS1-39	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAAGCAGAACCGGACAGAGCCCATAC					
IgkC1	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAAGCAGAACCGGACAGAGCCCATAC					
IgkC2	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAAGCAGAACCGGACAGAGCCCATAC					
IgkS1-34	241	251	261	271	281	291
IgkS1-35	AATATTGTAACCTTTTGTGCAAGTGTGACAGTACGCTTCGGTTGTGCGTACAAAGCACA					
IgkS1-36	AATATTGTAACCTTTTGTGCAAGTGTGACAGTACGCTTCGGTTGTGCGTACAAAGCACA					
IgkS1-37	AATATTGTAACCTTTTGTGCAAGTGTGACAGTACGCTTCGGTTGTGCGTACAAAGCACA					
IgkS1-38	AATATTGTAACCTTTTGTGCAAGTGTGACAGTACGCTTCGGTTGTGCGTACAAAGCACA					
IgkS1-39	AATATTGTAACCTTTTGTGCAAGTGTGACAGTACGCTTCGGTTGTGCGTACAAAGCACA					
IgkC1	AATATTGTAACCTTTTGTGCAAGTGTGACAGTACGCTTCGGTTGTGCGTACAAAGCACA					
IgkC2	AATATTGTAACCTTTTGTGCAAGTGTGACAGTACGCTTCGGTTGTGCGTACAAAGCACA					
IgkS1-34	301	311	321	331	341	351
IgkS1-35	CACGTAGACATTCTGACTTTTGGAAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC					
IgkS1-36	CACGTAGACATTCTGACTTTTGGAAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC					
IgkS1-37	CACGTAGACATTCTGACTTTTGGAAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC					
IgkS1-38	CACGTAGACATTCTGACTTTTGGAAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC					
IgkS1-39	CACGTAGACATTCTGACTTTTGGAAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC					
IgkC1	CACGTAGACATTCTGACTTTTGGAAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC					
IgkC2	CACGTAGACATTCTGACTTTTGGAAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC					
IgkS1-34	361	371	381			
IgkS1-35	ATCTGCAGTCAGAACCCCTAAGAATTC					
IgkS1-36	ATCTGCAGTCAGAACCCCTAAGAATTC					
IgkS1-37	ATCTGCAGTCAGAACCCCTAAGAATTC					
IgkS1-38	ATCTGCAGTCAGAACCCCTAAGAATTC					
IgkS1-39	ATCTGCAGTCAGAACCCCTAAGAATTC					
IgkC1	ATCTGCAGTCAGAACCCCTAAGAATTC					
IgkC2	ATCTGCAGTCAGAACCCCTAAGAATTC					

FIGURA 14

IgkC1	1	11	21	31	41	51
IgkS1-40	GGTACCGCCGCCACCATGGAGACGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkS1-41	GGTACCGCCGCCACCATGGAGACGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkS1-42	GGTACCGCCGCCACCATGGAGACTGACACTCTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkS1-43	GGTACCGCCGCCACCATGGAGACCGACACCCCTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkC2	GGTACCGCCGCCACCATGGAGACCGACACCCCTCCTGCTATGGGTGCTGCTGCTCTGGGTT					
IgkC1	61	71	81	91	101	111
IgkS1-40	CCAGGTTCCACTGGTGACGGATCCATGTCATGGAGATACACCTACATTCATGAATATATG					
IgkS1-41	CCAGGTTCCACGGGTGACGGATCCATGTCATGGAGATACACCTACGTTGTCATGAATATATG					
IgkS1-42	CCAGGTTCCACTGGTGACGGATCCATGTCATGGAGATACCTTACTTTGTCATGAATATATG					
IgkS1-43	CCAGGTTCCACCGGTGACGGATCCATGTCATGGAGATACCCCTACCTTTCATGAATATATG					
IgkC2	CCCGGCTCCACCGGCGACGGATCCATGCACGGCGACACCCCACTTGCACGAGTACATG					
IgkC1	121	131	141	151	161	171
IgkS1-40	TTAGATTGCAACCAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA					
IgkS1-41	TTAGATTGCAACCAGAGACGACGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA					
IgkS1-42	TTAGATTGCAACCAGAGACTACTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA					
IgkS1-43	TTAGATTGCAACCAGAGACCACCGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA					
IgkC2	CTGGACCTGCAGCCCGAGACCACCGGCTGTACGGCTACGGCCAGCTCAACGACAGCAGC					
IgkC1	181	191	201	211	221	231
IgkS1-40	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTTAC					
IgkS1-41	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTTAC					
IgkS1-42	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTTAC					
IgkS1-43	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTTAC					
IgkC2	GAGGAGGAGGACGAGATCGACGGCCCGCCGGCCAGGCCGAGCCGACCGCGCCCACTAC					
IgkC1	241	251	261	271	281	291
IgkS1-40	AATATTGTAACCTTTTGTGCAAGTGTGACTCTACGCTTCGGTTGTGCGGTACAAAGCACA					
IgkS1-41	AATATTGTAACCTTTTGTGCAAGTGTGACTCTACGCTTCGGTTGTGCGGTACAAAGCACA					
IgkS1-42	AATATTGTAACCTTTTGTGCAAGTGTGACTCTACTCTTCGGTTGTGCGGTACAAAGCACT					
IgkS1-43	AATATTGTAACCTTTTGTGCAAGTGTGACTCTACCTTCGGTTGTGCGGTACAAAGCACC					
IgkC2	AACATCGTGACCTTCTGTGCAAGTGTGACAGCACCTCTGCGCTCTGCGTGCAGAGCACC					
IgkC1	301	311	321	331	341	351
IgkS1-40	CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTTGTGTGCCCC					
IgkS1-41	CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTTGTGTGCCCC					
IgkS1-42	CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTTGTGTGCCCC					
IgkS1-43	CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACCTAGGAATTTGTGTGCCCC					
IgkC2	CACGTGACATCCGCACCTTGGAGGACCTGTGATGGGCACCTTGGGCATCGTGTGCCCC					
IgkC1	361	371	381			
IgkS1-40	ATCTGCTCTCAGAAGCCCTAAGAATTC					
IgkS1-41	ATCTGCTCTCAGAAGCCCTAAGAATTC					
IgkS1-42	ATCTGCTCTCAGAAGCCCTAAGAATTC					
IgkS1-43	ATCTGCTCTCAGAAGCCCTAAGAATTC					
IgkC2	ATCTGCTCCCAGAAGCCCTAAGAATTC					

FIGURA 15

	1	11	21	31	41	51
IgkC1	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGT					
IgkS1-44	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGT					
IgkS1-45	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGT					
IgkC2	GGTACCGCCGCCACCATGGAGACGACACCCTCCTGCTGTGGGTGCTGCTGCTCTGGGT					
	61	71	81	91	101	111
IgkC1	CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTCATGAATATATG					
IgkS1-44	CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTCATGAATATATG					
IgkS1-45	CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTCATGAATATATG					
IgkC2	CCCGGCTCCACGGGCGACGGATCCATGCACGGCGACACCCCACTGACAGGTACATG					
	121	131	141	151	161	171
IgkC1	TTAGATTGCAACCAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA					
IgkS1-44	TTAGATTGCAACCAGAGACAACCTGGTCTCTATGGTTATGGGCAATTAAATGACAGCTCA					
IgkS1-45	TTAGATTGCAACCAGAGACAACCTGGTCTCTACGGTTACGGGCAATTAAATGACAGCTCA					
IgkC2	CTGGACCTGCAGCCGAGACCACCGGCTGTACGGCTACGGCCAGCTCAACGACAGCAGC					
	181	191	201	211	221	231
IgkC1	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCCGGACAGAGCCCATTA					
IgkS1-44	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCCGGACAGAGCCCATTA					
IgkS1-45	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCCGGACAGAGCCCATTA					
IgkC2	GAGGAGGAGGACGAGATCGACGGCCCCCGCGCCAGGCCGAGCCGACCCGCGCCCACTAC					
	241	251	261	271	281	291
IgkC1	AATATTGTAACCTTTTGTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA					
IgkS1-44	AATATTGTAACCTTTTGTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA					
IgkS1-45	AATATTGTAACCTTTTGTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA					
IgkC2	AACATCGTGACCTTCTGCTGCAAGTGTGACAGCACCCTGCGCTCTGCGTGACAGCACC					
	301	311	321	331	341	351
IgkC1	CACGTAGACATTTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC					
IgkS1-44	CACGTAGACATTTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC					
IgkS1-45	CACGTAGACATTTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC					
IgkC2	CACGTGGACATCCGACCCCTGGAGGACCTGCTGATGGGCACCTGGGCATCGTGTGCCCC					
	361	371	381			
IgkC1	ATCTGCTCTCAGAAGCCCTAAGAATTC					
IgkS1-44	ATCTGCTCTCAGAAGCCCTAAGAATTC					
IgkS1-45	ATCTGCTCTCAGAAGCCCTAAGAATTC					
IgkC2	ATCTGCTCCAGAAGCCCTAAGAATTC					

FIGURA 16

IgkC1	1	11	21	31	41	51
IgkS1-46	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT					
IgkS1-47	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTGCTGCTGCTCTGGGTG					
IgkS1-48	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTA					
IgkS1-49	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTCTGCTGCTCTGGGTT					
IgkC2	GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTGCTGCTGCTCTGGGTC					
IgkC1	61	71	81	91	101	111
IgkS1-46	CCAGGTTCCACTGGTGACGGATCCATGTCATGGAGATACACCTACATTGCATGAATATATG					
IgkS1-47	CCAGGTTCCACTGGTGACGGATCCATGTCATGGAGATACACCTACATTGCATGAATATATG					
IgkS1-48	CCAGGTTCCACTGGTGACGGATCCATGTCATGGAGATACACCTACATTGCATGAATATATG					
IgkS1-49	CCAGGTTCCACTGGTGACGGATCCATGTCATGGAGATACACCTACATTGCATGAATATATG					
IgkC2	CCCGGCTCCACCGCGACGGATCCATGTCACGGCGACACCCACCCCTGCACGAGTACATG					
IgkC1	121	131	141	151	161	171
IgkS1-46	TTAGATTGCAACCAAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA					
IgkS1-47	TTAGATTGCAACCAAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA					
IgkS1-48	TTAGATTGCAACCAAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA					
IgkS1-49	TTAGATTGCAACCAAGAGACAACCTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA					
IgkC2	CTGACCTGCAAGCCGAGACACCGGCTGTACGGCTACGGCCAGCTCAACGACAGCAGC					
IgkC1	181	191	201	211	221	231
IgkS1-46	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATAC					
IgkS1-47	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATAC					
IgkS1-48	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATAC					
IgkS1-49	GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATAC					
IgkC2	GAGGAGGAGGACGAGATCGACGGCCCGCCGGCCAGGCCGAGCCGACCGCCGCCACTAC					
IgkC1	241	251	261	271	281	291
IgkS1-46	AATATTGTAACCTTTTGTGCAAGTGTGACTCTACGCTTCGGTTGTGCGGTACAAAGCACA					
IgkS1-47	AATATTGTAACCTTTTGTGCAAGTGTGACTCTACGCTTCGGTTGTGCGGTACAAAGCACA					
IgkS1-48	AATATTGTAACCTTTTGTGCAAGTGTGACTCTACGCTTCGGTTGTGCGGTACAAAGCACA					
IgkS1-49	AATATTGTAACCTTTTGTGCAAGTGTGACTCTACGCTTCGGTTGTGCGGTACAAAGCACA					
IgkC2	AACATCGTGACCTTCTGCTGCAAGTGCAGACAGCACCTGCGCCTCTGCGTCAGAGCACC					
IgkC1	301	311	321	331	341	351
IgkS1-46	CACGTAGACATTTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC					
IgkS1-47	CACGTAGACATTTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC					
IgkS1-48	CACGTAGACATTTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC					
IgkS1-49	CACGTAGACATTTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC					
IgkC2	CACGTGACATCCGCACCTGGAGGACCTGCTGATGGGCACCTGGGCATCGTGTGCCCC					
IgkC1	361	371	381			
IgkS1-46	ATCTGCTCTCAGAAGCCCTAAGAATTC					
IgkS1-47	ATCTGCTCTCAGAAGCCCTAAGAATTC					
IgkS1-48	ATCTGCTCTCAGAAGCCCTAAGAATTC					
IgkS1-49	ATCTGCTCTCAGAAGCCCTAAGAATTC					
IgkC2	ATCTGCTCCAGAAGCCCTAAGAATTC					

FIGURA 17

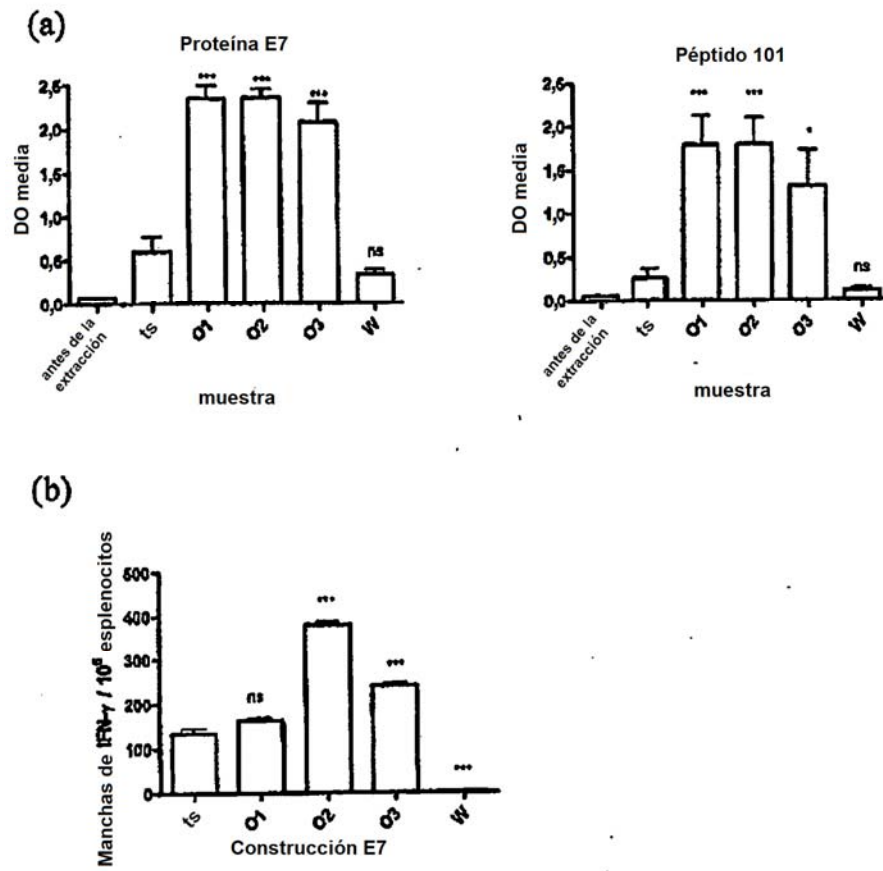


FIGURA 18

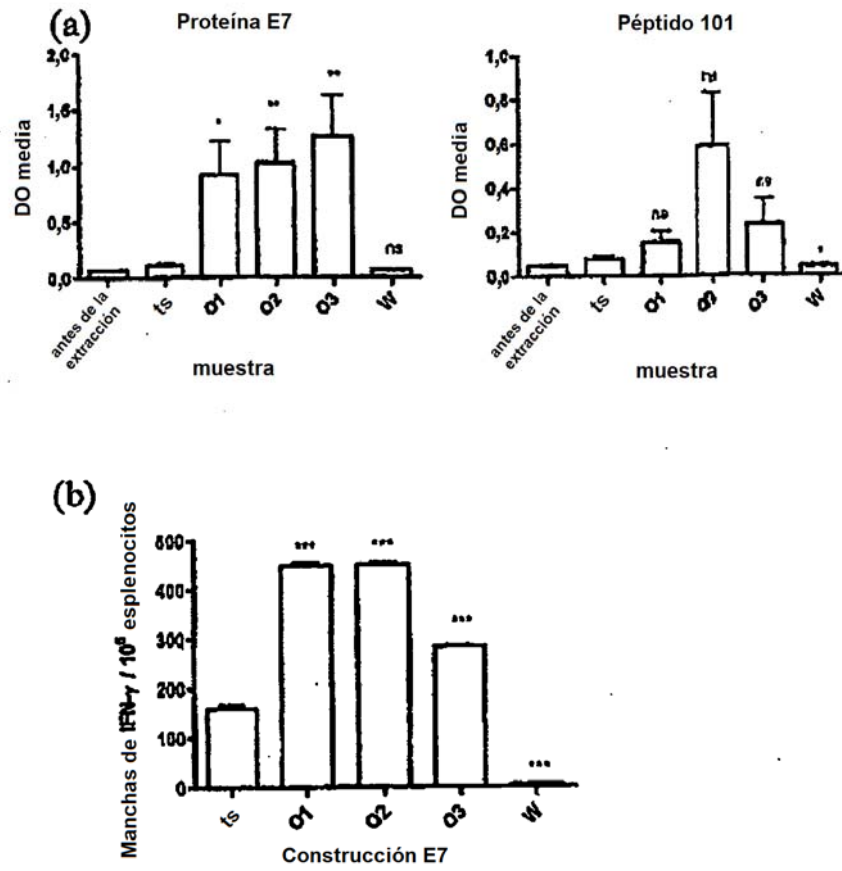


FIGURA 19

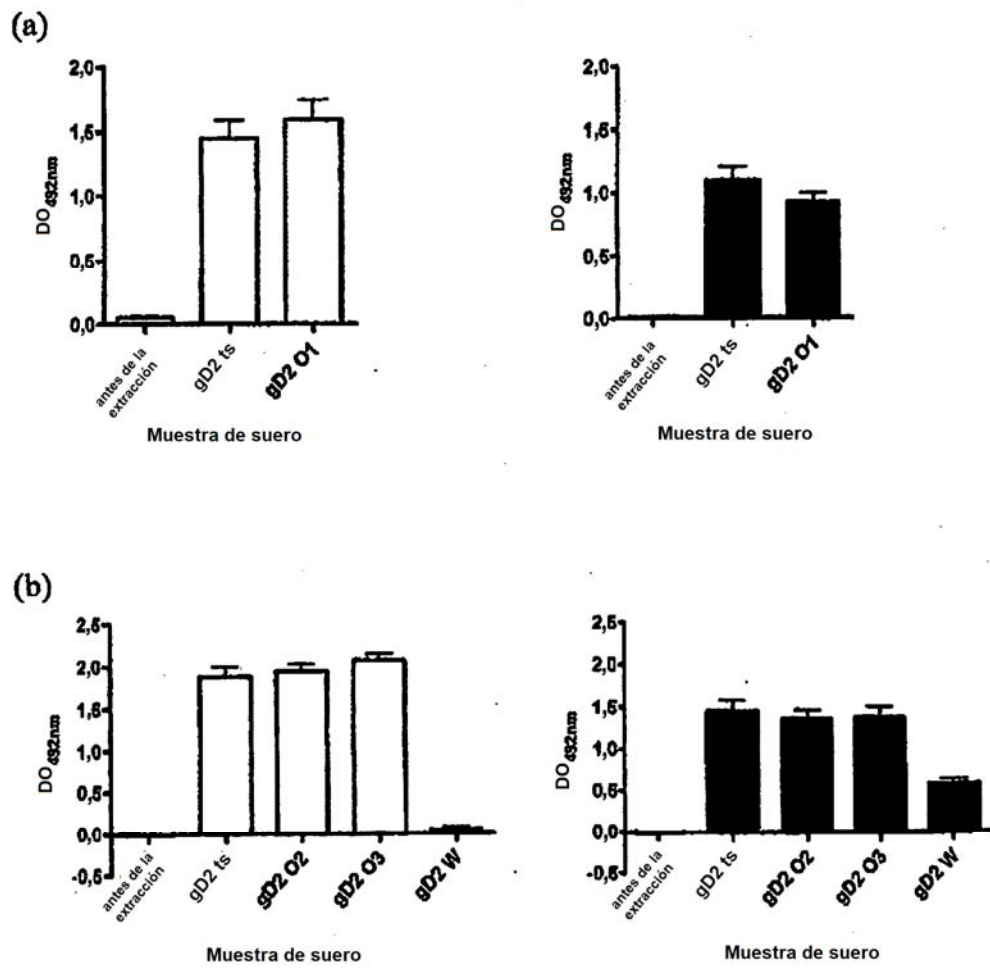


FIGURA 20