

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 873 083**

51 Int. Cl.:

C12N 7/00 (2006.01)

A61K 39/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **19.06.2015 PCT/EP2015/063815**

87 Fecha y número de publicación internacional: **23.12.2015 WO15193473**

96 Fecha de presentación y número de la solicitud europea: **19.06.2015 E 15731884 (1)**

97 Fecha y número de publicación de la concesión europea: **10.03.2021 EP 3158059**

54 Título: **Método para la generación rápida de un virus de ARN atenuado**

30 Prioridad:

20.06.2014 EP 14305956

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

03.11.2021

73 Titular/es:

UNIVERSITÉ D'AIX-MARSEILLE (50.0%)

58, BD Charles Livon

13007 Marseille, FR y

INSERM (INSTITUT NATIONAL DE LA SANTÉ ET DE LA RECHERCHE MÉDICALE) (50.0%)

72 Inventor/es:

NOUGAIREDE, ANTOINE;

DE FABRITUS, LAURIANE;

AUBRY, FABIEN;

DE LAMBALLERIE, XAVIER y

GOULD, ERNEST ANDREW

74 Agente/Representante:

VEIGA SERRANO, Mikel

ES 2 873 083 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

Método para la generación rápida de un virus de ARN atenuado

5 **Sector de la técnica**

La presente invención se refiere a un método para la generación *in vitro* rápida de virus ARN positivos monocatenarios atenuados de receptores humanos o animales que pueden usarse como vacunas para proteger contra infecciones y enfermedades víricas.

10

Estado de la técnica

Muchas enfermedades infecciosas emergentes están provocadas por virus de ARN monocatenario. Los principales brotes de dengue, encefalitis del Nilo occidental, fiebre Chikungunya y fiebre del Valle del Rift que se han producido en las últimas décadas, cada uno con un impacto significativo en la salud humana, destacan la urgente necesidad de conocer los factores que permiten que estos virus invadan nuevos territorios o se adapten a nuevas especies hospedadoras o vectores.

15

Estos eventos a menudo se perciben como señales de advertencia de una posible pandemia. En el caso de una pandemia, comprender los factores que dan forma a la adaptabilidad de estos agentes infecciosos en rápida evolución y nuestra capacidad para desarrollar rápidamente una vacuna serán los pasos críticos para controlar la propagación de la enfermedad.

20

De hecho, hasta la fecha, la vacunación sigue siendo el mejor método para reducir la mortalidad y la morbilidad de los seres humanos provocadas por estos virus. En particular, las vacunas vivas atenuadas tienen un gran éxito debido a la estimulación de diferentes brazos de la respuesta inmunitaria del hospedador. Estas vacunas vivas atenuadas son variantes de virus naturales derivadas al pasar virus en hospedadores anormales. Sin embargo, la preparación de una vacuna viva atenuada padece muchos inconvenientes, especialmente porque su preparación se basa en un método empírico y que requiere mucho tiempo. Por lo tanto, actualmente se necesita mucho tiempo para desarrollar una vacuna útil que se pueda administrar a los seres humanos.

25

30

Existe por lo tanto una necesidad insatisfecha de un enfoque de generación de virus atenuados, que no tiene posibilidad de reversión y que proporciona un método rápido, eficiente, rentable y seguro de fabricación de una vacuna candidata.

35

La presente invención satisface esta necesidad proporcionando un enfoque sistemático para diseñar futuras vacunas candidatas que esencialmente no tienen posibilidad de reversión. Este método es ampliamente aplicable a un amplio intervalo de virus y proporciona un enfoque eficaz para producir una amplia diversidad de vacunas antivíricas.

40 **Objeto de la invención**

La presente invención aprovecha el poder de la mutagénesis para producir un virus de ARN atenuado en un período muy corto, es decir, tan pronto como se conozca la secuencia completa del virus diana y se pueda producir un genoma infeccioso.

45

Debido a que hay más codones que aminoácidos, el código genético es necesariamente redundante. Los diferentes codones que codifican el mismo aminoácido se conocen como codones sinónimos. A menudo se supone que los cambios en la secuencia de ADN de una proteína entre dos codones sinónimos no tienen ningún efecto y, por lo tanto, se denominan mutación sinónima. Sin embargo, aunque los codones sinónimos codifican los mismos aminoácidos, los inventores han demostrado que la sustitución de sinónimos en grandes regiones del genoma vírico da como resultado la atenuación eficaz del virus (Nougairede et al, Random Codon Re-encoding Induces Stable Reduction of Replicative Fitness of Chikungunya Virus in primate and Mosquito Cells, PLOS Pathogens, 2013). De manera más precisa, los inventores descubrieron que el reemplazo de los codones de nucleótidos nativos del genoma de un virus de ARN por codones de nucleótidos sinónimos disminuye la capacidad de replicación del virus, atenuando de esta manera dicho virus.

50

55

Los inventores también desarrollaron un enfoque novedoso para generar virus de ARN que no requieren clonación y propagación de un ADNc de longitud completa en una bacteria. Esta tecnología se basa en la observación de que la superposición de fragmentos de ADN bicatenario, cada uno cubriendo una porción del genoma vírico, permiten espontáneamente la recombinación y síntesis de una copia de ADN del genoma vírico completo después de la transfección.

60

Combinando estos dos enfoques, los inventores desarrollaron un método para generar directamente un virus atenuado, que tiene varias ventajas para el desarrollo de candidatos a vacunas, incluyendo la posibilidad de obtener candidatos a vacuna en un período muy corto, tan pronto como se conozca la secuencia completa del patógeno dirigido y pueda producirse un genoma infeccioso. El método de la invención es por lo tanto extremadamente útil para generar,

65

en días, una vacuna viva atenuada dirigida contra un patógeno novedoso para el que no se dispone de tratamiento o vacuna.

En consecuencia, **en un primer aspecto**, la invención se refiere a un método para generar *in vitro* un virus de ARN positivo monocatenario atenuado de un receptor humano o animal que comprende las siguientes etapas:

etapa I) recodificar el genoma vírico de un virus ARN infeccioso sustituyendo aleatoriamente una parte de los codones de nucleótidos de todo el genoma vírico de dicho virus ARN infeccioso por otro codón de nucleótidos que codifica el mismo aminoácido, con la condición de que:

- i) el número y la posición de los codones de nucleótidos raros presentes en dicho genoma vírico no se modifiquen, siendo dichos codones de nucleótidos raros CGU, CGC, CGA, CGG, UCG, CCG, GCG y ACG; y
- ii) las regiones de dicho genoma vírico que están implicadas con la estructura secundaria del ARN no se modifiquen.

etapa II) generar un virus de ARN atenuado mediante:

sub-etapa II.a) introducción de un promotor de la ARN polimerasa dependiente de ADN en la posición 5' y opcionalmente un terminador y una secuencia de poliadenilación de ARN en la posición 3' del genoma vírico recodificado como se obtiene en la etapa I);

sub-etapa II.b) amplificación del genoma vírico recodificado como se prepara en la sub-etapa a) incluyendo dicho promotor y opcionalmente dicho terminador y secuencia de poliadenilación de ARN, en al menos 2, preferentemente al menos 3, 4, 5 o 6 fragmentos de ADNc solapantes;

sub-etapa II.c) transfección de dichos fragmentos de ADNc en una célula hospedadora,

sub-etapa II.d) incubación de dicha célula hospedadora de la sub-etapa c); y

sub-etapa II.e) recuperación del virus de ARN infeccioso de dicha célula hospedadora incubada.

En un segundo aspecto, la divulgación pertenece a una composición farmacéutica que comprende un virus de ARN atenuado obtenido de acuerdo con el método desvelado en el presente documento.

En un tercer aspecto, la divulgación se refiere al uso del método desvelado en el presente documento para desarrollar una vacuna viva atenuada, o al uso del virus de ARN atenuado obtenido de acuerdo con el método desvelado en el presente documento como una vacuna viva atenuada.

Descripción detallada de la invención

En un primer aspecto, la invención se refiere a un método para generar *in vitro* un virus de ARN positivo monocatenario atenuado de un receptor humano o animal que comprende las siguientes etapas:

etapa I) re-codificar el genoma vírico de un virus ARN infeccioso sustituyendo aleatoriamente una parte de los codones de nucleótidos de todo el genoma vírico de dicho virus ARN infeccioso por otro codón de nucleótidos que codifica el mismo aminoácido, con la condición de que:

- i) el número y la posición de los codones de nucleótidos raros presentes en dicho genoma vírico no se modifiquen, siendo dichos codones de nucleótidos raros CGU, CGC, CGA, CGG, UCG, CCG, GCG y ACG; y
- ii) las regiones de dicho genoma vírico que están implicadas con la estructura secundaria del ARN no se modifiquen.

etapa II) generar un virus de ARN atenuado mediante:

sub-etapa II.a) introducción de un promotor de la ARN polimerasa dependiente de ADN en la posición 5' y opcionalmente un terminador y una secuencia de poliadenilación de ARN en la posición 3' del genoma vírico recodificado como se obtiene en la etapa I);

sub-etapa II.b) amplificación del genoma vírico recodificado como se prepara en la sub-etapa a) incluyendo dicho promotor y opcionalmente dicho terminador y secuencia de poliadenilación de ARN, en al menos 2, preferentemente al menos 3, 4, 5 o 6 fragmentos de ADNc solapantes;

sub-etapa II.c) transfección de dichos fragmentos de ADNc en una célula hospedadora;

sub-etapa II.d) incubación de dicha célula hospedadora de la sub-etapa c); y

sub-etapa II.e) recuperación del virus de ARN infeccioso de dicha célula hospedadora incubada. Basado en sus exhaustivas investigaciones, los inventores desarrollaron una estrategia muy prometedora para generar

directamente un virus atenuado mediante la recodificación a gran escala. En consecuencia, la divulgación proporciona un virus atenuado, que comprende un genoma vírico modificado que contiene sustituciones de nucleótidos diseñadas en múltiples ubicaciones en el genoma, en donde las sustituciones introducen una pluralidad de codones sinónimos en el genoma. El término "**virus atenuado**", como se usa en el presente documento, se refiere a un virus con virulencia comprometida en el receptor previsto, por ejemplo, receptor humano o animal. Más específicamente, un virus atenuado tiene una capacidad disminuida o debilitada de producir enfermedad mientras retiene la capacidad de estimular una respuesta inmunitaria similar a la del virus de tipo silvestre.

Esta novedosa estrategia representa una ruta significativamente mejorada para el desarrollo de vacunas. De hecho, la re-codificación dirigida al sitio, asociada a ninguna modificación de las secuencias de aminoácidos, alivia la probabilidad de novedosas propiedades fenotípicas y, por lo tanto, proporciona beneficios para el desarrollo genérico de vacunas vivas atenuadas, incluyendo costes reducidos e inducción de una dosis única de inmunidad a largo plazo.

Etapa de re-codificación a gran escala

El método de la invención comprende una primera etapa I) de mutagénesis, también denominada "**recodificación a gran escala**" en lo siguiente. Como se usa en el presente documento, las expresiones "**método de re-codificación**" o "**método de recodificación a gran escala**" se refieren a una etapa de recodificación del genoma vírico de un virus de ARN, preferentemente una región de dicho genoma vírico, sustituyendo aleatoriamente una parte de los codones de nucleótidos de todo el genoma vírico de dicho virus ARN infeccioso por otro codón de nucleótidos que codifica el mismo aminoácido, con la condición de que:

- i) el número y la posición de los codones de nucleótidos raros presentes en dicho genoma vírico no se modifiquen, siendo dichos codones de nucleótidos raros CGU, CGC, CGA, CGG, UCG, CCG, GCG y ACG; y
- ii) las regiones de dicho genoma vírico que están implicadas con la estructura secundaria del ARN no se modifiquen.

Preferentemente, la **etapa I)** es una etapa de recodificación del genoma vírico de un virus de ARN infeccioso mediante la sustitución aleatoria de aproximadamente el 1 % a aproximadamente el 20 %, preferentemente de aproximadamente el 1 % a aproximadamente el 17 %, preferentemente de aproximadamente el 1 % a aproximadamente el 15 %, preferentemente de aproximadamente el 1 a aproximadamente el 10 %, preferentemente de aproximadamente el 3 a aproximadamente el 8 %, preferentemente de aproximadamente el 3 % a aproximadamente el 5 % de los codones de nucleótidos de todo el genoma vírico de dicho virus ARN infeccioso por otro codón de nucleótidos que codifica el mismo aminoácido, con la condición de que:

- i) el número y la posición de los codones de nucleótidos raros presentes en dicho genoma vírico no se modifiquen, siendo dichos codones de nucleótidos raros CGU, CGC, CGA, CGG, UCG, CCG, GCG y ACG; y
- ii) las regiones de dicho genoma vírico que están implicadas con la estructura secundaria del ARN no se modifiquen.

El método de re-codificación modifica de esta manera la composición de ácidos nucleicos de grandes regiones codificantes del genoma vírico del virus ARN sin modificar las proteínas codificadas mediante la introducción de un gran número de mutaciones sinónimas.

El material de partida de la etapa I) es preferentemente un virus de ARN infeccioso. Preferentemente, el genoma del virus se re-codifica de tal manera que de aproximadamente el 1 % a aproximadamente el 20 % de los codones de nucleótidos están sustituidos por diferentes codones de nucleótidos que codifican el mismo aminoácido. Esto es posible gracias al sesgo de uso de codones.

Como se usa en el presente documento, las expresiones "**codones de nucleótidos sinónimos**" o "**codones sinónimos**" se refieren a dos o más codones de nucleótidos que codifican el mismo aminoácido. De hecho, la mayoría de los aminoácidos están codificados por más de un codón. Los codones sinónimos son codones que codifican el mismo aminoácido.

Como se usa en el presente documento, las expresiones "**mutación sinónima**" o "**sustitución de sinónimos**" se refieren a la sustitución de un codón de nucleótidos por otro codón de nucleótidos que codifica el mismo aminoácido, es decir, un codón sinónimo. Los inventores han demostrado que las sustituciones sinónimos reducen la capacidad de replicación de un virus. Adicionalmente, la introducción de codones sinónimos en un genoma de un virus limita su capacidad para mutar o usar la recombinación para volverse virulento. Es de destacar que para la obtención de un virus ARN atenuado, que todavía podrían usarse como vacunas vivas atenuadas, solo del 1 % al 20 %, preferentemente del 1 % al 10 % de los codones de nucleótidos del genoma vírico se recodifican aleatoriamente. En el contexto de la invención, las mutaciones sinónimas se introducen mediante mutagénesis dirigida al sitio. Preferentemente, dichas mutaciones se insertan mediante mutagénesis en casete.

Mientras que la mayoría de los aminoácidos pueden estar codificados por varios codones diferentes, no todos los

codones se usan con la misma frecuencia, algunos codones son codones **"raros"**. Como se usa en el presente documento, un codón **"raro"** es uno de al menos dos codones sinónimos que codifican un aminoácido particular que está presente en un ARNm a una frecuencia significativamente más baja que el codón usado con más frecuencia para ese aminoácido. Normalmente, dichos codones raros son CGU, CGC, CGA, CGG, UCG, CCG, GCG y ACG.

5 El método diseñado por los inventores se basa en la observación de que dichos codones raros deben permanecer inalterados para controlar eficazmente la atenuación vírica.

10 Adicionalmente, los inventores llegaron a la conclusión de que las regiones del genoma vírico implicadas en la estructura secundaria del ARN no se modificarán para controlar eficazmente la atenuación vírica. En consecuencia, dichas regiones no deben re-codificarse.

15 Como se usa en el presente documento, la expresión **"regiones del genoma vírico implicadas con la estructura secundaria del ARN"** se refiere a regiones conservadas del genoma del virus, que contienen estructuras de ARN funcionalmente activas, también conocida como "estructura secundaria de ARN". Se demuestra que dichas estructuras de ARN son importantes durante las diversas etapas del ciclo de vida vírico. El experto en la materia determinaría fácilmente las regiones implicadas con una estructura secundaria de ARN significativa, que suelen estar bien conservadas en la filogenia evolutiva.

20 Básicamente, la etapa I) de recodificar el genoma de un virus de ARN infeccioso comprende la siguiente etapa:

- identificación de codones en múltiples ubicaciones dentro de porciones no reguladoras del genoma vírico, cuyos codones pueden reemplazarse por codones sinónimos, no siendo dichos codones CGU, CGC, CGA, CGG, UCG, CCG, GCG y ACG;
- 25 - seleccionar un codón sinónimo para sustituir entre aproximadamente el 1 y aproximadamente el 20 % de los codones de nucleótidos identificados; y
- sustituir un codón sinónimo para cada uno de los codones identificados, preferentemente sobre la base de la tabla 1 a continuación.

30 Preferentemente, la etapa I) se realiza:

- determinando la secuencia de aminoácidos codificada por todo el genoma vírico del virus de ARN infeccioso y determinar cada codón de nucleótido que codifica cada aminoácido; y
- sustituyendo aproximadamente el 1 % a aproximadamente el 20 %, preferentemente de aproximadamente el 1 % a aproximadamente el 17 %, preferentemente de aproximadamente el 1 % a aproximadamente el 15 %, preferentemente de aproximadamente el 1 a aproximadamente el 10 %, preferentemente de aproximadamente el 3 a aproximadamente el 8 %, preferentemente de aproximadamente el 3 % a aproximadamente el 5 % del codón de nucleótidos del genoma vírico que codifica un aminoácido de la tabla 1, por un codón de nucleótido diferente que codifica el mismo aminoácido como se especifica en la tabla 1:

40 Tabla 1: Mutación sinónima de acuerdo con la invención

Aminoácido	Codón de nucleótido
Ala, A	GCU, GCC, GCA
Arg/R	AGA, AGG
Asn/N	AAU, AAC
Asp/D	GAU, GAC
Cys/C	UGU, UGC
Gln/Q	CAA, CAG
Glu/E	GAA, GAG
Gly/G	GGU, GGC, GGA
His/H	CAU, CAC
Ile/I	AUU, AUC, AUA
Leu/L	UUA, UUG, CUU, CUC, CUA, CUG
Lys/K	AAA, AAG

(continuación)

Aminoácido	Codón de nucleótido
Phe/F	UUU, UUC
Pro/P	CCU, CCC, CCA
Ser/S	UCU, UCC, UCA, AGU, AGC
Thr/T	ACU, ACC, ACA
Tyr/Y	UAU, UAC
Val/V	GUU, GUC, GUA, GUG

5 Los virus atenuados, obtenidos de acuerdo con la invención, tienen la notable propiedad de no modificarse a nivel proteico. De hecho, dichos virus atenuados corresponden a virus que fueron modificados genéticamente mediante sustituciones sinónimas, sólo a nivel nucleico.

10 La atenuación vírica puede confirmarse de formas que son bien conocidas por un experto en la materia. Los ejemplos no limitantes incluyen ensayos en placa, mediciones de crecimiento y reducción de la morbilidad o letalidad en animales de prueba.

Más específicamente, los inventores han demostrado que la recodificación a gran escala tiene un impacto en la aptitud de replicación del virus de ARN diana, atenuando de esta manera dicho virus.

15 **"Aptitud replicativa"** se define como la capacidad de replicación/adaptabilidad de un organismo en un entorno dado. La aptitud replicativa de un virus, o un virus atenuado obtenido de acuerdo con el método descrito en el presente documento, puede medirse *in cellulo*, por ejemplo, por medios que implican competiciones entre dos o más cepas víricas en cultivo de tejidos. Normalmente, la aptitud replicativa puede determinarse una vez que se recupera el virus, en diversos tipos de células tales como células de primates no humanos o células de mosquitos en el caso de los arbovirus. Normalmente, el virus recodificado a evaluar y un virus de tipo silvestre se pasan en diversos tipos de
20 células, tales como células de primates no humanos (Vero) o de mosquito (C6/36). La cinética replicativa de cada virus de pases se determina usando técnicas conocidas como el análisis de la tasa de crecimiento vírico, basado en el análisis de los valores de TCID50.

25 Los inventores han demostrado que la etapa de recodificación aleatoria disminuye la aptitud replicativa del virus tanto en primates (como en células de artrópodos en el caso de los arbovirus). La disminución de la aptitud de replicación se correlacionó directamente con el grado de recodificación. Estos resultados corroboran que la recodificación de codones reduce profundamente el título infeccioso del virus liberado mientras que el número de partículas víricas permanece estable.

30 El virus de ARN atenuado que el método pretende generar (también denominado **"virus diana"** en el presente documento) es una versión atenuada de un virus de ARN positivo monocatenario de un receptor humano o animal. Más preferentemente dicho virus se selecciona del grupo que consiste en flavivirus, alfavirus y enterovirus.

35 Una lista no limitante de flavivirus comprende el virus del dengue (DENV), Virus de la fiebre amarilla (YFV), Encefalitis de San Luis (SLEV), Virus de la encefalitis japonesa (JEV), Encefalitis del valle de Murray (MVEV), Virus del Nilo Occidental (VNO), Rocio (ROCV), Virus de la encefalitis transmitida por garrapatas (TBEV), Fiebre hemorrágica de Omsk (OMSKV), Enfermedad de Kyasanr Forrest (KFDV), Powassan (POWV). Preferentemente, dicho flavivirus se selecciona del grupo que consiste en:

- 40
- Virus de la encefalitis japonesa (JEV); tal como una cepa de genotipo I (JEV I) o la cepa de genotipo III (JEV III),
 - Virus del Nilo Occidental (VNO), tal como una cepa de genotipo 2;
 - Virus del dengue (DENV), tal como una cepa de serotipo 4;
 - Virus de la fiebre amarilla (YFV), tal como una cepa de tipo silvestre de América del Sur; y
 - Virus de la encefalitis transmitida por garrapatas (TBEV), tal como una cepa de subtipo del Lejano Oriente.
- 45

Más preferentemente, dicho flavivirus es el virus de la encefalitis transmitida por garrapatas

Una lista no limitante de alfavirus comprende el virus Chikungunya (CHIK), encefalitis equina del este (EEE), virus de la encefalitis equina occidental, virus de la encefalitis equina venezolana (VEE), virus Mayaro (MAYO), virus O'nyong'nyong (ONN), virus Sindbis, virus del bosque de Semliki, virus del bosque de Barmah, virus del río Ross, virus Una, virus Tonate. Preferentemente, dicho alfavirus es el virus Chikungunya.

50

Una lista no limitante de enterovirus comprende Coxsackie, ecovirus, poliovirus y rinovirus. Preferentemente, dicho enterovirus es Coxsackie, más preferiblemente virus Coxsackie B.

En una realización preferida, el virus diana es el virus Chikungunya. En esta realización específica, la etapa de recodificación se realiza en tres regiones del genoma vírico, en concreto:

- 5
- la región que codifica la proteína no estructural nsP1;
 - la región que codifica la proteína no estructural nsP4; y
 - la región que se solapa con las proteínas estructurales E2 y E1.

10 Normalmente, las mutaciones se introducen gracias a la mutagénesis de casete, también denominado "casetes recodificados". Normalmente, se usa un casete recodificado de aproximadamente 1300 pb a aproximadamente 1500 pb para cada región.

En esta realización, el genoma vírico de Chikungunya se modifica como sigue:

- 15
- la región que codifica la proteína no estructural nsP1, en la posición 242-1543 (nt) del genoma completo está mutado por un casete recodificado de 1302 nt, como se representa en SEQ ID NO: 63;
 - la región que codifica la proteína no estructural nsP4, en la posición 6026-7435 (nt) del genoma completo está mutado por un casete recodificado de 1410 nt, como se representa en SEQ ID NO: 64; y
 - la región que se solapa con las proteínas estructurales E2 y E1, en la posición 9526-11022 (nt) del genoma
- 20 completo está mutado por un casete recodificado de 1500 nt, como se representa en SEQ ID NO: 65.

Normalmente, cada uno de los casetes recodificados introduce de 200 a 400 mutaciones sinónimas, preferentemente de aproximadamente 250 a aproximadamente 320, preferentemente de aproximadamente 266 a aproximadamente 320. Preferentemente, el casete recodificado como se representa en SEQ ID NO: 63, SEQ ID NO: 64 y SEQ ID NO: 25 65 introducen respectivamente 264, 298 y 320 mutaciones sinónimas.

En otra realización, el virus diana es el virus de la encefalitis transmitida por garrapatas (TBEV). En esta realización específica, la etapa de recodificación se realiza en la región genómica NS5 del virus, que codifica la proteína no estructural NS5. Normalmente se usa un casete recodificado de aproximadamente 1400 pb. En esta realización, el 30 genoma vírico de TBEV se modifica para introducir de aproximadamente 200 a aproximadamente 350 mutaciones sinónimas, preferentemente de aproximadamente 200 a aproximadamente 300, preferentemente de aproximadamente 225 a aproximadamente 300, preferentemente de aproximadamente 225 a aproximadamente 275, preferentemente de aproximadamente 225.

35 Preferentemente, el genoma vírico del virus de la encefalitis transmitida por garrapatas está mutado por un casete recodificado de 1412 nt, representado en SEQ ID NO: 66. Dicho casete recodificado introduce 273 mutaciones.

En aún otra realización, el virus diana es el virus de la encefalitis japonesa (JEV), preferiblemente una cepa de genotipo 1. En esta realización específica, la etapa de recodificación se realiza en una región grande del genoma vírico, típicamente en casi todo el marco de lectura abierto (ORF) completo, desde el comienzo de PrM hasta el final de la 40 región genómica NS5.

Preferentemente, el genoma vírico de JEV se modifica para introducir de aproximadamente 163 a aproximadamente 658, preferentemente de aproximadamente 163 a aproximadamente 658 mutaciones.

45 Preferentemente, el genoma viral del virus de la encefalitis japonesa está mutado por al menos un casete recodificado seleccionado del grupo que consiste en:

- 50
- casete recodificado Ia, como se representa en SEQ ID NO: 67;
 - casete recodificado Ib, como se representa en SEQ ID NO: 68;
 - casete recodificado IIa, como se representa en SEQ ID NO: 69;
 - casete recodificado IIb, como se representa en SEQ ID NO: 70;
 - casete recodificado IIIa, como se representa en SEQ ID NO: 71; y
 - casete recodificado IIIb, como se representa en SEQ ID NO: 72.

55 Más preferentemente, el genoma vírico del virus de la encefalitis japonesa se muta mediante una combinación de casetes recodificados de la siguiente manera:

- 60
- combinación de casetes recodificados Ia e IIa; o
 - combinación de casetes recodificados Ia e IIIa; o
 - combinación de casetes recodificados IIa e IIIa.

Generación directa de un virus de ARN atenuado

65 El método de la invención comprende una segunda etapa II) de generación directa de un virus de ARN atenuado.

Los inventores desarrollaron un enfoque novedoso para generar directamente un virus de ARN atenuado, partiendo del genoma vírico recodificado aleatoriamente de dicho virus. Los inventores evidenciaron que los fragmentos de ADNc solapantes, abarcando cada uno una porción del genoma de un virus de ARN, puede dar lugar a un virus sin el uso de un ADNc de longitud completa o un plásmido o un vector que comprenda tal ADNc de longitud completa. En consecuencia, los inventores aclararon que los fragmentos de ADN de doble cadena solapantes, abarcando cada uno una porción del genoma vírico atenuado, permiten espontáneamente la recombinación y síntesis de una copia de ADN del genoma vírico completo *in cellulo*.

Dicho método es altamente ventajoso, especialmente porque exonera de:

- construir un ADNc de longitud completa, que abarca el genoma vírico recodificado completo; y/o
- el uso de un plásmido o un vector que comprende dicho ADNc de longitud completa; y/o
- la necesidad de reconstruir el ADNc de longitud completa o el genoma vírico atenuado completo antes de la transfección en una célula hospedadora; y/o
- modificar el genoma vírico atenuado tal como incorporando recombinación de origen no natural o sitios de enzimas de restricción; y/o
- uso de virus auxiliares u otra proteína vírica.

Esta etapa II) específica de la invención, también denominada "**Amplicones subgenómicos infecciosos**" o "**ISA**", es por lo tanto un procedimiento muy sencillo capaz de acelerar la producción de virus de ARN atenuados en unos días, con perfecto control de las secuencias víricas y partiendo de un genoma vírico recodificado.

La estrategia se basa en la producción de varios fragmentos de ADNc, abarcando cada uno un fragmento del genoma vírico recodificado. El montaje de la construcción no se realiza *in vitro* mediante ensamblaje de Gibson o clonación de extensión de polimerasa circular antes de la transfección, sino a través de un proceso de recombinación que tiene lugar directamente *in cellulo*.

Como se usa en el presente documento, la expresión "**generación de virus de ARN atenuados**" se refiere a la producción de un virus de ARN, en una forma genéticamente modificada, es decir, en una forma recodificada de acuerdo con el método de la invención.

Como se usa en el presente documento, la expresión "**sitio de recombinación que no se produce de forma natural**" se refiere a secuencias que permiten la recombinación específica de sitio que pueden ejemplificarse mediante los sistemas de recombinación Cre-Lox o FLP-FRT. El sitio de la enzima de restricción se refiere a secuencias que permiten el corte específico del sitio de ADN bicatenario mediante enzimas de restricción que pueden ejemplificarse mediante las endonucleasas NotI o AluI.

La etapa II) del método de la invención comprende una **sub-etapa II.a)** de introducir un promotor de la ARN polimerasa dependiente de ADN en la posición 5' de todo el genoma de un virus de ARN. Opcionalmente, dicha sub-etapa II.a) comprende además la introducción de un terminador y una secuencia de poliadenilación de ARN en la posición 3' del genoma completo de un virus de ARN.

Es de destacar que cuando el virus diana es un virus poliadenilado, tal como flavivirus, la sub-etapa II.a) es una etapa de introducción de un promotor de la ARN polimerasa dependiente de ADN en la posición 5' y un terminador y una secuencia de poliadenilación de ARN en la posición 3' del genoma completo de un virus de ARN.

Incluyendo, en el extremo 5' del primer fragmento de ADNc, un promotor de la ARN polimerasa dependiente de ADN, y en el extremo 3' del último fragmento de ADNc una secuencia de ribozima y una secuencia señal para la poliadenilación del ARN, el fragmento de ADNc se transcribe como un genoma atenuado de ARN de longitud completa con extremos 5' y 3' auténticos.

Preferentemente, dicho promotor de la ARN polimerasa dependiente de ADN en la posición 5' es el promotor del citomegalovirus humano (pCMV), como se representa en SEQ ID NO:1. Preferentemente, dicha secuencia de poliadenilación de terminador y ARN es, respectivamente, la ribozima de la hepatitis delta y la señal de poliadenilación del virus de simio 40 (HDR/SV40pA). La secuencia de HDR/SV40pA se representa en SEQ ID NO: 2.

En consecuencia, la sub-etapa a) proporciona el genoma vírico recodificado completo del virus de ARN, flanqueado respectivamente en 5' y 3' por el promotor del citomegalovirus humano (pCMV) (SEQ ID NO: 1) y la ribozima delta de la hepatitis seguida por la señal de poliadenilación del virus de simio 40 (HDR/SV40pA) (SEQ ID NO:2)

La etapa II) del método de la invención comprende una **sub-etapa II.b)** de amplificación de todo el genoma vírico recodificado en varios fragmentos de ADNc solapantes. En la sub-etapa II.b), el genoma vírico completo corresponde al genoma vírico completo preparado en la etapa a), es decir, que incluye dicho promotor y opcionalmente dicho terminador y secuencia de poliadenilación de ARN.

Como se usa en el presente documento, la expresión "**fragmentos de ADNc solapantes**", "**fragmentos de ADNc**",

también designados **"amplicones"** o **"fragmentos subgenómicos de ADN"** o **"amplicones subgenómicos"** son fragmentos de ADN bicatenario que abarcan solo una parte del genoma vírico recodificado de un virus de ARN.

Tales fragmentos corresponden a **"fragmentos subgenómicos"**.

Los inventores aclararon que, cuando dichos fragmentos se transfectan dentro de una célula, sorprendentemente se recombinan espontáneamente *in cellulo* para reconstituir todo el genoma vírico recodificado. Dicha recombinación se produce incluso si el genoma vírico no está modificado genéticamente para incorporar sitios de recombinación adicionales y que no se producen de forma natural.

Los fragmentos de ADNc de acuerdo con la invención comprenden:

- fragmentos de ADN obtenidos por amplificación, por ejemplo por PCR; así como
- fragmentos de ADN obtenidos *de novo*.

Preferentemente, dichos fragmentos de ADNc no son infecciosos.

Como se usa en el presente documento, la expresión **"ADNc de longitud completa"**, se refiere a un ADN que comprende todo el genoma vírico de un virus en una sola pieza, preferentemente el genoma vírico recodificado completo.

Como se usa en el presente documento, la expresión **"Fragmento de ADNc que abarca una parte del genoma vírico recodificado completo"**, se refiere a un fragmento de ADN que comprende una porción del genoma vírico recodificado completo. Normalmente, los fragmentos de ADNc de acuerdo con la invención se recombinan espontáneamente tras la transfección en células para constituir una copia de ADN de todo el genoma vírico recodificado, flanqueado en el extremo 5' por un promotor de la ARN polimerasa dependiente de ADN, y en el extremo 3' por una secuencia de terminación y una secuencia señal para la poliadenilación del ARN. Esta construcción se transcribe como un genoma recodificado por ARN de longitud completa con extremos 5' y 3' auténticos por la maquinaria celular. Por el contrario, un **"ADNc de longitud completa que abarca el genoma vírico completo"** es un solo ADNc que codifica la totalidad del genoma vírico, preferentemente la totalidad del genoma vírico recodificado.

Preferentemente, la etapa II.b) del método de la invención permite la producción de 2 a 15 fragmentos de ADNc solapantes, preferentemente de 3, 4, 5 o 6 fragmentos de ADNc solapantes. Normalmente, dichos fragmentos de ADNc son de aproximadamente 2 kb a aproximadamente 6 kb, preferentemente de aproximadamente 4 kb y cada fragmento de ADNc tiene aproximadamente 70 a aproximadamente 100 pb de regiones solapantes.

Preferentemente, dichos fragmentos de ADNc solapantes de la etapa II.b) son:

- fragmentos de clon infeccioso no amplificados por PCR;
- fragmentos de clon infeccioso amplificados por PCR;
- fragmentos de clon no infeccioso no amplificados por PCR;
- fragmentos de clon no infeccioso amplificados por PCR;
- fragmentos sintetizados *de novo* no amplificados por PCR;
- fragmentos sintetizados *de novo* amplificados por PCR; y
- fragmentos obtenidos por PCR con transcripción inversa del genoma vírico.

La etapa II) del método *in vitro* de la invención comprende una **sub-etapa II.c)** de transfección de dichos fragmentos de ADNc en una célula hospedadora.

Como se usa en el presente documento, el término **"transfección"** se refiere a la introducción de ácidos nucleicos (ya sea ADN o ARN) en células eucariotas o procariotas. Una célula que ha absorbido el ácido nucleico exógeno se denomina "célula hospedadora" o "célula transfectada". La transfección puede lograrse mediante una diversidad de medios conocidos en la técnica que incluyen la co-precipitación de ADN con fosfato de calcio, transfección mediada por DEAE-dextrano, transfección mediada por polibreno, electroporación, microinyección, fusión de liposomas, lipofección, fusión de protoplastos, infección retroviral y biolística.

Preferentemente, la célula hospedadora de la sub-etapa c) es una célula permisiva, que permite la recuperación de un virus infeccioso. Normalmente, las células permisivas empleadas en el método de la presente invención son células que, tras la transfección con los fragmentos de ADNc, son capaces de realizar un ciclo de replicación completo del virus, incluyendo la producción de partículas víricas. Preferentemente, dicha célula hospedadora se selecciona del grupo que consiste en líneas celulares SW13, BHK-21, HEK 293 y Vero.

En una realización preferida, la sub-etapa II.c) es una etapa de transfección directa de los fragmentos de ADNc obtenidos en la sub-etapa II.b) como tal, y la sub-etapa II.c) se produce directamente después de la sub-etapa II.b). En esta realización específica, los fragmentos de ADNc como tales se transfectan a las células hospedadoras. Dichos fragmentos se recombinan espontáneamente *in cellulo* en una copia de ADN del genoma vírico recodificado completo

flanqueado en el extremo 5' por un promotor de la ARN polimerasa dependiente de ADN, y en el extremo 3' por una secuencia de terminación y una secuencia señal para la poliadenilación del ARN. Como se ha mencionado previamente, el método de la invención supera un prejuicio técnico ya que exonera de transfectar un ADNc de longitud completa, abarcando el genoma vírico completo, como tal. Aparte, el método está exento de usar un plásmido o un vector que comprende dicho ADNc de longitud completa como tal y/o la necesidad de reconstruir el ADNc completo o el genoma vírico completo antes de la transfección en una célula hospedadora.

Por el contrario, el método se basa en la transfección de los fragmentos de ADNc solapantes, comprendiendo cada uno una porción del genoma vírico recodificado. La transfección de fragmentos de ADN bicatenario solapantes, que abarcan el genoma completo de un virus de ARN, en células permisivas permite espontáneamente la recombinación y síntesis de una copia de ADN del genoma vírico completo *in cellulo*.

En una realización alternativa, la sub-etapa II.c) es una etapa de transfección de plásmidos, cada uno de los cuales comprende un fragmento de ADNc obtenido en la sub-etapa II.b), en donde cada fragmento de ADNc se incorpora en plásmidos o vectores individuales y separados.

En esta realización, cada fragmento de ADNc se incorpora en plásmidos o vectores individuales y separados. Cada plásmido o vector comprende un solo fragmento de ADNc. En esta realización, el genoma vírico recodificado completo se reconstituye después de la transfección.

En una realización, el método de la invención comprende una etapa II.b') adicional después de la sub-etapa b) y antes de la sub-etapa c) de purificación de los fragmentos de ADNc solapantes. Dicha purificación puede realizarse mediante cualquier técnica conocida, preferentemente a través de una columna de cromatografía.

La etapa II) del método de la invención comprende una **sub-etapa II.d)** de incubación de las células hospedadoras, que preferentemente dura de 3 a 9 días. Durante dicha etapa de incubación, los fragmentos de ADNc transfectados se recombinan espontáneamente en las células hospedadoras para constituir una copia de ADN de todo el genoma vírico recodificado, flanqueado en el extremo 5' por un promotor de la ARN polimerasa dependiente de ADN, y en el extremo 3' por una secuencia de terminación y una secuencia señal para la poliadenilación del ARN. Esta construcción se transcribe como un genoma de ARN de longitud completa con extremos 5' y 3' auténticos por la maquinaria celular.

En consecuencia, el producto rescatado en la sub-etapa II.d) es un virus de ARN atenuado.

Composición farmacéutica

El virus de ARN atenuado obtenido de acuerdo con el método desvelado en el presente documento puede estar comprendido en una composición farmacéutica.

Todos los datos técnicos desvelados anteriormente son aplicables aquí.

Dichas composiciones farmacéuticas que comprenden virus atenuados son adecuadas para la inmunización.

Preferentemente, la administración de una composición farmacéutica tal puede ser por diversas vías parenterales tales como vía subcutánea, intravenosa, intradérmica, intramuscular, intraperitoneal, intranasal, oral o transdérmica. La administración parenteral puede lograrse mediante inyección en bolo o por perfusión gradual en el tiempo.

Las preparaciones para administración parenteral incluyen soluciones, suspensiones y/o emulsiones acuosas o no acuosas estériles, que pueden contener agentes auxiliares o excipientes conocidos en la técnica. Algunos ejemplos de disolventes no acuosos son propilenglicol, polietilenglicol, aceites vegetales tales como aceite de oliva y ésteres orgánicos inyectables tales como oleato de etilo. Pueden usarse vehículos o vendajes oclusivos para aumentar la permeabilidad de la piel y mejorar la absorción de antígenos. Las formas de dosificación líquidas para administración oral pueden comprender una solución de liposomas que contienen la forma de dosificación líquida. Las formas adecuadas para suspender liposomas incluyen emulsiones, suspensiones, soluciones, jarabes y elixires que contienen diluyentes inertes usados comúnmente en la técnica, tales como agua purificada. Además de los diluyentes inertes, tales composiciones también pueden incluir adyuvantes, agentes humectantes, agentes emulsionantes y de suspensión, o agentes edulcorantes, aromatizantes o perfumantes.

Cuando se usa una composición de la presente divulgación para la administración a un individuo, puede comprender además sales, tampones, adyuvantes u otras sustancias que sean deseables para mejorar la efectividad de la composición. Para vacunas, pueden usarse adyuvantes, sustancias que pueden aumentar una respuesta inmunitaria específica. Normalmente, el adyuvante y la composición se mezclan antes de la presentación al sistema inmunitario, o se presentan por separado, pero en el mismo sitio del organismo que se inmuniza.

La administración de la composición puede ser con un propósito "profiláctico" o "terapéutico". Cuando se proporcionan de manera profiláctica, las composiciones de la divulgación que son vacunas se proporcionan antes de que se manifieste cualquier síntoma o signo clínico de una infección por patógenos. La administración profiláctica de la

composición sirve para prevenir o atenuar cualquier infección posterior. La divulgación pertenece a la composición farmacéutica como se desvela en el presente documento para su uso para prevenir una infección por virus de ARN en un sujeto.

- 5 Cuando se proporciona terapéuticamente, una vacuna vírica atenuada se proporciona tras la detección de un síntoma o signo clínico de una infección real. La administración terapéutica sirve para atenuar cualquier infección real. La divulgación se refiere a la composición farmacéutica desvelada en el presente documento para su uso para tratar una infección por virus de ARN en un sujeto.
- 10 Por lo tanto, puede proporcionarse una composición de vacuna atenuada de la presente divulgación antes del inicio de la infección (para prevenir o atenuar una infección anticipada) o después del inicio de una infección real.

Diseño de futura vacuna candidata

- 15 La divulgación se refiere al uso del método desvelado en el presente documento para desarrollar una vacuna viva atenuada.

Todos los datos técnicos desvelados anteriormente son aplicables aquí.

- 20 Se ha demostrado que la etapa de recodificación de codones a gran escala de la invención es un poderoso método de atenuación que tiene varias ventajas para el desarrollo de vacunas, incluyendo la posibilidad de obtener cepas de vacunas potenciales en un período muy corto tan pronto como se conozca la secuencia completa del patógeno diana y pueda producirse un genoma infeccioso. El método de la invención es por lo tanto extremadamente útil para generar, en días, una vacuna viva atenuada dirigida contra un patógeno novedoso para el que no se dispone de tratamiento o vacuna.

Adicionalmente, los inventores han demostrado que el método de la invención es ventajoso en varios aspectos al diseñar futuras vacunas candidatas, en concreto:

- 30 (i) la reversión al tipo silvestre es intrínsecamente más difícil, dado el alto número de mutaciones producidas;
- (ii) dado que la reducción de la aptitud replicativa disminuye con el grado de recodificación, el método abre la puerta para ajustar con precisión la reducción de la aptitud mediante la modulación de la longitud de las regiones recodificadas y el número de mutaciones sinónimas introducidas;
- 35 (iii) el uso de una combinación de varias regiones recodificadas ubicadas en todo el genoma vírico previene la reversión fenotípica completa debido a la recombinación entre virus TS y recodificados: la modificación de secuencia a gran escala hace que la recombinación sea intrínsecamente más difícil y, en el caso de la recombinación, la parte del genoma que representa la cepa recodificada probablemente aún portaría algunas mutaciones asociadas a la reducción de la aptitud.

- 40 Todos los datos técnicos desvelados anteriormente son aplicables aquí.

Los siguientes ejemplos se dan con el fin de ilustrar diversas realizaciones de la invención y no pretenden limitar la presente invención de ninguna manera.

45 Descripción de las figuras

Figura 1: Estrategia universal para rescatar virus de ARN positivos monocatenarios.

- El genoma vírico completo, representado esquemáticamente en la figura (genoma flavivírico), flanqueado respectivamente en 5' y 3' por el promotor del citomegalovirus humano (pCMV) y la ribozima delta de la hepatitis seguida de la señal de poliadenilación del virus de simio 40 (HDR/SV40pA), se amplificó mediante PCR en 3 fragmentos de ADNc solapantes. La transfección de productos de PCR en células permisivas permitió la recuperación de virus después de 3 a 9 días. Las flechas azules horizontales representan los cebadores usados para generar los 3 fragmentos de ADNc solapantes.

Figura 2: Representación esquemática de los virus recodificados CHIKV.

- De arriba a abajo: Barra de escala de nucleótidos; representación esquemática del genoma completo de CHIKV con regiones codificantes (rectángulos grises), no codificantes (rectángulos negros) y la cola poliA. Las regiones recodificadas se representan en gris oscuro.

Figura 3: Representación esquemática de los 9 JEV recodificados diferentes obtenidos con el método ISA.

- Cada rectángulo representa un fragmento. Los rectángulos morados se usan cuando no se introdujeron mutaciones (TS). Los rectángulos azules (bajo nivel de recodificación) y verdes (alto nivel de recodificación) se usan para los fragmentos recodificados (el valor representa el número de mutaciones sinónimas).

Figura 4: Aptitud replicativa de los JEV TS y recodificados.

- La aptitud replicativa *in cellulo* de los JEV recodificados se midió usando células humanas. Los resultados muestran una disminución de la aptitud replicativa de acuerdo con el nivel de recodificación, el tamaño de la región de recodificación y la posición genómica del fragmento o fragmentos recodificados.

Ejemplos

EJEMPLO 1: Generar un virus de ARN en unos días

- 5 El Ejemplo 1 ilustra el método ISA que permite la producción de un virus de ARN en días.

La siguiente ilustración de ISA se basa en genomas víricos que no se modificaron previamente, es decir, genoma vírico que no pasó por una etapa de recodificación.

10 MÉTODOS

Células, virus, clones infecciosos y anticuerpos

- 15 Se cultivaron células de riñón de cría de hámster (BHK-21) a 37 °C con CO₂ al 5 % en un medio esencial mínimo (Life Technologies) con suero bovino fetal inactivado por calor al 7 % (FBS; Life Technologies) y Penicilina/Estreptomina al 1 % (PS; 5000 U/ml y 5000 µg/ml; Life Technologies). Las células de riñón embrionario humano 293 (HEK-293) y las células de riñón de mono verde africano (VeroE6) se cultivaron a 37 °C con CO₂ al 5 % en el mismo medio que las células BHK-21 suplementadas con 1 % de aminoácidos no esenciales (Life technologies). Se cultivaron células de carcinoma suprarrenal humano (SW13) a 37 °C con CO₂ al 5 % en medio RPMI 1640 (Life Technologies) con FBS al 10 % y PS al 1 %. La cepa JEV de genotipo I JEV_CNS769_Laos_2009 (KC196115) se aisló en junio de 2009 del líquido cefalorraquídeo de un paciente en Laos16; YFV cepa BOL 88/1999 (KF907504), aislada en 2009 de un suero humano, fue amablemente proporcionada por el Centro Nacional de Enfermedades Tropicales (CENETROP), Santa-Cruz, Bolivia; Cepa DENV-4 Dak HD 34 460 (KF907503), aislada de un suero humano, fue proporcionada amablemente por Robert B Tesh del Center for Biodefense and Emerging Infectious Diseases - Sealy Center for Vaccine Development (University of Texas Medical Branch, Galveston, Texas, USA); el clon infeccioso del genotipo III de JEV derivado de la cepa rp9 (DQ648597) fue amablemente proporcionado por Yi-Ling Lin del Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwán; el clon infeccioso de WNV derivó de la cepa Ouganda 1937 (M12294); el clon infeccioso de TBEV derivó de la cepa Oshima 5.10 (AB062063); el clon infeccioso de CV-B3 derivó de la cepa 2679 (KJ489414). Se usaron un suero inmune específico de JEV (obtenido después de la vacunación contra JEV) y anticuerpos específicos de DENV monoclonales¹⁷ para realizar ensayos de inmunofluorescencia directa.

Preparación de fragmentos de ADNc

- 35 El genoma completo flanqueado respectivamente en 5' y 3' por el promotor del citomegalovirus humano (pCMV) (SEQ ID NO: 1) y la ribozima delta de la hepatitis seguida por la señal de poliadenilación del virus de simio 40 (HDR/SV40pA) (SEQ ID NO:2) se amplificó mediante PCR en tres fragmentos de ADN solapantes de aproximadamente 4,8 kb, 3,0 kb y 4,3 kb (4,8 kb, 2,9 kb y 5,2 kb para CHIKV, 4,8 kb, 4,1 kb y 3,4 kb para TBEV y 2,9 kb, 2,8 kb y 2,7 kb para CV-B3) (véase la Tabla 1 a continuación).

- 40 Para WNV, TBEV, JEV III y CHIKV, los fragmentos de ADN se obtuvieron mediante PCR usando clones infecciosos (para JEV III, se corrigió una mutación usando PCR de fusión).

- 45 Para JEV I (todos los fragmentos de ADN), DENV-4 (primer y tercer fragmento) e YFV (primer y tercer fragmento), Los fragmentos de ADN se sintetizaron de novo (Genscript) y se amplificaron mediante PCR. Los amplicones se produjeron usando el kit Platinum PCR SuperMix High Fidelity (Life Technologies).

- 50 La mezcla (volumen final: 50 µl) consistía en 45 µl de supermezcla, 2 µl de molde de ADN a 1 ng/µl (clon infeccioso o fragmento de ADN sintetizado) y 200 nM de cada cebador. Para DENV-4 y YFV, el segundo fragmento de ADN se obtuvo mediante RT-PCR a partir de sobrenadantes celulares clarificados. Los ácidos nucleicos se extrajeron usando el EZ1 Virus Mini Kit v2 en el EZ1 Biorobot (ambos de Qiagen) de acuerdo con las instrucciones del fabricante y se amplificaron con el kit Superscript III One-Step RT-PCR Platinum Taq Hifi (Life Technologies). La mezcla (volumen final: 50 µl) contenía 25 µl de mezcla de reacción, 2 µl de extracto de ácido nucleico, 100 nM de cada cebador, 1 µl de mezcla de enzimas y 20 µl de agua libre de nucleasas. Los ensayos se realizaron en un termociclador de gradiente convencional Biometra T-professional con las siguientes condiciones: 94 °C durante 2 min seguido de 40 ciclos de 94 °C durante 15 s, 64 °C durante 30 s, 68 °C durante 5 min y una etapa preliminar de 50 °C durante 30 min para la RT-PCR. El tamaño de los productos de PCR se verificó mediante electroforesis en gel y se purificó usando un kit Amicon Ultra - 0,5 ml (Millipore) de acuerdo con las instrucciones del fabricante. Cuando se usó ADN plasmídico como molde, la retirada completa del molde se aseguró mediante una etapa de digestión con la enzima de restricción Dpn1 (New England Biolabs) antes de la transfección. Para controlar la eficiencia de esta etapa adicional, los inventores transflectaron (véase a continuación), como control, sólo dos fragmentos de ADNc (el primero y el segundo, 1 µg final). Estos controles no produjeron ningún virus infeccioso.

Tabla 2: Cebadores usados para obtener fragmentos de ADNc

Virus	Fragmento de ADNc	Cebador directo	Posición	SEQ ID	Cebador inverso	Posición	SEQ ID
JEV I	I	CACCCAACTGATCTTCAGCATCT	-	3	GAAGAATGATTTCTGTAAGTGTCCAG	4054-4078	4
	II	CGTTGCCATGCCAATCTTAGCG	4002-4023	5	GGTGCTTGCCTCCACCAA	6983-7004	6
	III	CAAAATGAGTATGGAATGCTGGAAAA	6932-6956	7	CTCAGGGTCAATGCCAGCGCTT	-	8
JEV II	I	GCCCACCGGAAGGAGCTGAC	-	9	CAGAGAGCAAAATCCCTATGACGA	4078-4100	10
	II	CGTCACCATGCCAGTCTTAGCG	4001-4022	11	GCTTGGCAATCCAGTCAGTCCT	7004-7025	12
	III	CAAAACGAGTACGGAATGCTAGAAA	6931-6954	13	CTCATGTTTGACAGCTTATCATCG	-	14
WNV	I	TCAATATTGGCCATTAGCCATATTAT		15	TGGATTGAACACTCCTGTAGACGC	4135-4158	16
	II	TGGTTGGAGTTGGAAGCCTCATC	4052-4074	17	GACCATGCCGTGGCCGGCC	7016-7034	18
	III	TGGACAAGACCAAGAATGACATTG	6920-6943	19	GTTACAAATAAAGCAATAGCATCACA	-	20
TBEV	I	CAGGGTTATTGTCTCATGAGCGGA	-	21	GCCACGCCCCAGGAAGAGCATGA	4033-4054	22
	II	GGGCCCTCTGGAATGGGAGA	3892-3913	23	CAACCCAGGCTTGTACCACATCTT	8003-8026	24
	III	GGGTGAGTCTGGACCTTGGGA	7886-7907	25	CCTAGGAATTTACAAAAATAAAGCATTTT	-	26
YFV	I	CACCCAACTGATCTTCAGCATCT	-	27	GCATGGAAGTGTCTTTGAGTTCT	4071-4094	28
	II	GACTTGCAACGATGCTCTTTTGCA	4020-4043	29	GAGAGAGCATCGTCACAATGCC	7040-7061	30
	III	GATTCATCCAGCACCCGCCACC	6964-6984	31	CTCAGGGTCAATGCCAGCGCTT	-	32
DENV-4	I	GAATAAGGGCGACACGGAAATGT		33	TGAAGACAGCTTGTCTGCACAA	-	34
	II	GATCATGGCTTGGAGGACCATTAT	3980-4003	35	GCTACTGCATAGAGCGTCCATG	6949-6970	36
	III	TTTACCAGGTAAAAACAGAAACCAC	6892-6916	37	CTCAGGGTCAATGCCAGCGCTT		38
fragmentos I 6 JEV	I	CACCCAACTGATCTTCAGCATCT		39	CATGGAACCATTCCTTATGGACT	1635-1657	40
	II	ACTGGATTGTGAACCAAGGAGTG	1560-1582	41	GAAGAATGATTTCTGTAAGTGTCCAG	4054-4078	42
	III	CGTTGCCATGCCAATCTTAGCG	4002-4023	43	AATATAACCCCGAGCGGCGGATG	5511-5532	44
	IV	ATGTCACCAAAACAGGGTGCCCAA	5440-5462	45	GGTGTCTGGCTCCCTCCACCAA	6983-7004	46
	V	CAAAATGAGTATGGAATGCTGGAAAA	6932-6956	47	GCGCCGTGCTCCATTGATTCTG	8950-8971	48
	VI	GGCTGTGGGCACATTTGTACCG	8843-8864	49	CTCAGGGTCAATGCCAGCGCTT	-	50

(continuación)

Virus	Fragmento de ADNc	Cebador directo	Posición	SEQ ID	Cebador inverso	Posición	SEQ ID
CHIKV	I	CACCCAACTGATCTTCAGCATCT		51	CTGCTCGGGTGACCTGTCCTA	4050-4070	52
	II	TGAGATGTTTTTCCCTATTCAGCAACT	3961-3986	53	AACAATGTGTTGACGAACAGAGTTA	6966-6990	54
	III	CTCCCTGCTGGACTTGATAGAG	6859-6880	55	CTCAGGGTCAATGCCAGCGCTT	-	56
CV-B3	I	CACCCAACTGATCTTCAGCATCT		57	CCACACAACATGCGTACCAAGCA	2184-2206	58
	II	CAGGCGCTGGCGCTCCGACA	2148-2167	59	GTCATGGTTATACTCTCTGAACA	4970-4994	60
	III	GACAGGAGGACACAAAGTCAGAT	4921-4943	61	CTCAGGGTCAATGCCAGCGCTT	-	62

Transfección celular

Se incubó 1 µg final de una mezcla equimolar de todos los fragmentos de ADNc amplificados por PCR o 1 µg de clon infeccioso de CV-B3 con 12 µl de Lipofectamine 2000 (Life Technologies) en 600 µl de medio Opti-MEM (Life Technologies). De acuerdo con las instrucciones del fabricante, la mezcla se añadió a un matraz de cultivo de 12,5 cm² de células subconfluentes que contenía 1 ml de medio sin antibióticos. Después de 4 horas de incubación, se retiró el sobrenadante celular, las células se lavaron dos veces (HBSS; Life Technologies) y se añadieron 3 ml de medio fresco. El sobrenadante celular se recogió cuando se observó un efecto citopático bruto (CPE) (3-9 días dependiendo del tipo de célula y la velocidad de crecimiento del virus) o 9 días después de la transfección para virus no citopáticos, se clarificó por centrifugación, se tomaron alícuotas y se almacenó a -80 °C. A continuación, cada virus se pasó cuatro veces utilizando el mismo tipo de célula excepto para DENV-4 e YFV para los que se usaron respectivamente VeroE6 y HEK-293. Los pases se realizaron inoculando 333 µl de sobrenadante celular clarificado en células en un matraz de cultivo de 12,5 cm² que contenía 666 µl de medio: después de 2 horas de incubación, las células se lavaron dos veces (HBSS) y se añadieron 3 ml de medio fresco. El sobrenadante celular se recogió después de 2-6 días, se clarificó por centrifugación, se tomaron alícuotas y se almacenó a -80 °C. Se usaron sobrenadantes celulares clarificados (reservas de virus) para realizar la cuantificación del ARN vírico, ensayo TCID₅₀, ensayo de inmunofluorescencia directa y secuenciación del genoma completo.

Ensayos de PCR y RT-PCR en tiempo real

Para evaluar la producción de virus infecciosos y asegurarse de que la detección positiva no fue el resultado de la contaminación por ADNc, se cuantificó el ARN vírico y se comparó con la cantidad de ADNc detectado utilizando el kit Access RT-PCR Core Reagent (Promega) con o sin la transcriptasa inversa. El ARN se extrajo usando el kit EZ1 mini virus 2.0 y el EZ1 Biorobot (ambos de Qiagen) de acuerdo con las instrucciones del fabricante. La mezcla (volumen final: 25 µl) contenía una cantidad convencional de tampón de reacción AMV/Tfl 5X, 0,5 µM de cada cebador, 0,5 µl de mezcla de dNTP, 0,5 mM de MgSO₄, 0,5 µl de transcriptasa inversa AMV (solo para RT-PCR), 0,5 µl de ADN polimerasa de Tfl, 15,5 µl de agua libre de nucleasas y 2 µl de ácidos nucleicos extraídos. Los ensayos se realizaron usando el sistema de detección de PCR en tiempo real CFX96 Touch™ (Biorad) con las siguientes condiciones: 50 °C durante 15 min, 95 °C durante 2 min, seguido de 45 ciclos a 95 °C durante 15 s, 60 °C durante 40 s. La recopilación de datos se produjo durante la etapa de 60 °C. La diferencia entre los valores de umbral de ciclo (ct) obtenidos mediante ensayos de PCR en tiempo real y RT-PCR en tiempo real se ha usado para evaluar la producción de ARN vírico. Adicionalmente, la cantidad de ARN vírico expresada como límite de detección de dosis (unidad arbitraria; AU) se calculó a partir de curvas patrón (se usaron como estándar ácidos nucleicos de sobrenadantes celulares de virus cultivados; se combinaron cinco extractos de ácido nucleico y se almacenaron alícuotas de 10 µl a -80 °C).

Ensayo de dosis 50 infecciosa de cultivo de tejidos (TCID₅₀)

Para cada determinación, se inoculó un cultivo en placa de 96 pocillos que contenía 20.000 células BHK-21 en 100 µl de medio por pocillo (añadido justo antes de la inoculación) con 50 µl de diluciones seriadas 10 veces de sobrenadantes de cultivo celular clarificado: cada fila incluía 6 pocillos de la misma dilución y dos controles negativos. Las placas se incubaron durante 7 días y se leyó la ausencia o presencia de CPE en cada pocillo. La determinación del DICT₅₀/ml se realizó mediante el método de Reed y Muench¹⁸.

Ensayo de inmunofluorescencia directa (dIFA)

Se realizaron IFA directas usando frascos de cultivo de 12,5 cm² de células SW13 para JEV I y JEV III, y las células VeroE6 se infectaron respectivamente 2 y 6 días antes de usar el sobrenadante celular clarificado (véase anteriormente: pasaje de virus). El sobrenadante se retiró y las células se lavaron dos veces (HBSS; Invitrogen), se tripsinizaron, se recogieron y se diluyeron (1/5) con medio fresco. Después de la citocentrifugación de 150 µl de esta suspensión celular (3 min, 900 rpm; Cytospin, Thermo Scientific), los portaobjetos se secaron, se sumergieron 20 min en acetona fría para la fijación, se secaron, se incubaron 30 min a 37 °C con suero inmunitario específico de JEV adecuadamente diluido (véase anteriormente) o anticuerpos monoclonales específicos de DENV, se lavaron dos veces con PBS, se lavaron una vez con agua destilada, se secaron, se incubaron durante 30 min a 37 °C con el anticuerpo secundario conjugado con FITC adecuadamente diluido y se contratiñó con azul de Evans, se lavó dos veces con PBS, se lavó una vez con agua destilada, se secó, se montó y se leyó usando un microscopio de fluorescencia.

Análisis de secuencia del genoma de longitud completa

La secuenciación completa del genoma se realizó usando el Ion PGM Sequencer¹⁹ (Life Technologies) y los análisis se realizaron con el software CLC Genomics Workbench 6. Los sobrenadantes de virus se clarificaron primero y se trataron con Benzonasa nucleasa HC > 99 % (Novagen) a 37 °C durante la noche. Después de la extracción de ARN (no se usó vehículo de ARN; véase anteriormente) usando el kit EZ1 mini virus 2.0 y el EZ1 Biorobot (ambos de Qiagen), la amplificación aleatoria de ácidos nucleicos se realizó como se describió previamente²⁰. El ADN amplificado se analizó usando el secuenciador Ion PGM de acuerdo con las instrucciones del fabricante. Las lecturas obtenidas se recortaron: primero usando la puntuación de calidad, después retirando los cebadores usados durante la amplificación aleatoria y finalmente en las extremidades 5' y 3' retirando sistemáticamente 6 nucleótidos. Solo se

usan lecturas con una longitud superior a 29 nucleótidos y se asignan a la secuencia del genoma original usada como una referencia. Las frecuencias de mutación (proporción de genomas víricos con la mutación) para cada posición se calcularon simplemente como el número de lecturas con una mutación en comparación con la referencia dividido por el número total de lecturas en ese sitio.

5

RESULTADOS

Los inventores desarrollaron una genética inversa sencilla y versátil que facilita la recuperación de virus de ARN infecciosos a partir de material de ADN genómico sin requerir clonación, propagación de ADNc en bacterias o transcripción de ARN *in vitro*. Su hipótesis de trabajo era que la transfección de fragmentos de ADN bicatenario solapantes, que abarcan el genoma completo de un virus de ARN, en células susceptibles permitiría espontáneamente la recombinación y síntesis de una copia de ADN del genoma vírico completo. Incluyendo en el extremo 5' del primer (5') fragmento de ADN, un promotor de ARN polimerasas dependientes de ADN, y en el extremo 3' del último (3') fragmento de ADN una secuencia de ribozima y una secuencia señal para la poliadenilación del ARN, los inventores anticiparon que esta copia de ADN genómico se transcribiría como un genoma de ARN de longitud completa con extremos 5' y 3' auténticos que se exportarían eficientemente fuera del núcleo (en el caso de un virus que se replica en el compartimento citoplásmico).

Los inventores probaron primero esta hipótesis con 6 flavivirus (es decir, virus envueltos transmitidos por artrópodos con un genoma de ARN monocatenario de polaridad positiva que se replican en el citoplasma de las células infectadas) que representan los principales linajes evolutivos flavivíricos: dos virus de la encefalitis japonesa (JEV; genotipo I (JEV I) y genotipo III (JEV III)), un virus del Nilo Occidental de genotipo 2 (WNV), un virus del dengue de serotipo 4 (DENV-4), una cepa de tipo salvaje del virus de la fiebre amarilla (YFV) y un virus de la encefalitis transmitida por garrapatas del subtipo del Lejano Oriente (TBEV) (**Tabla 2**).

25

Se amplificaron genomas completos mediante PCR en 3 fragmentos de ADN de aproximadamente 4 kb, cada uno con regiones solapantes de 70-100 pb. El primer y el último fragmentos estaban flanqueados respectivamente en 5' y 3' por el promotor del citomegalovirus humano (pCMV) y la ribozima delta de la hepatitis seguido de la señal de poliadenilación del virus de simio 40 (HDR/SV40pA) (Figura 1). Los productos de PCR se purificaron en columna y se transfeció 1 µg de una mezcla equimolar de todos los fragmentos en líneas celulares SW13 y/o BHK-21, los cuales, aseguran la recuperación eficiente de los genomas infecciosos flavivíricos. Los medios sobrenadantes de células de estos cultivos infecciosos se pasaron en serie cuatro veces usando los mismos tipos de células, permitiendo el aislamiento de JEV I, JEV III, TBEV y WNV. Para virus más exigentes, el aislamiento podría lograrse mediante el paso en células permisivas adicionales (por ejemplo, DENV-4: células VeroE6; YFV: células HEK-293). Se demostró la replicación del virus después de cuatro pases en serie para cada virus usando una combinación de los siguientes criterios:

- (i) producción de genomas víricos en medio sobrenadante celular usando métodos de RT-PCR en tiempo real,
- (ii) producción de partículas infecciosas en medio sobrenadante celular usando ensayos TCID50,
- (iii) detección de efecto citopático (CPE),
- (iv) detección de antígenos víricos mediante ensayos de inmunofluorescencia directa, y
- (v) secuenciación completa del genoma vírico usando el método de secuenciación de próxima generación (NGS).

40

La robustez, la flexibilidad y la versatilidad de los métodos fueron desafiadas adicionalmente como sigue. En primer lugar, los inventores redujeron el tamaño y aumentaron el número de fragmentos solapantes combinados para la transfección. Esto se ejemplificó en el caso de JEV I, para el cual el método ISA generó virus infecciosos, cuando se usan hasta 6 amplicones solapantes de aproximadamente 2 kb. En segundo lugar, aplicaron el método ISA a virus con un genoma de ARN monocatenario de polaridad positiva que pertenecen a diferentes familias: virus de Chikungunya (CHIKV, un virus con envuelta, familia Togaviridae) y Coxsackievirus B3 (CV-B3, un virus no envuelto, familia Picornaviridae). De nuevo, los virus infecciosos pudieron aislarse después de la transfección y cuatro pases en células HEK-293 (CHIKV) o células BGM (CV-B3) (Tabla 2 a continuación).

Adicionalmente, los inventores usaron como control el CV-B3 obtenido tras la transfección de un genoma infeccioso que lleva plásmidos y obtuvieron resultados similares en términos de infectividad y datos de secuencia (**Tabla 3**).

55

Virus	Cepa	Origen del material usado para producir amplicones subgenómicos			Línea celular usada para transfección	Línea celular usada durante pasajes	RT-PCR en tiempo real (UA)	Log10 TCD ₅₀ /ml	CFE	dIFA	dN/dS (todas las mutaciones)	dN/dS (mutaciones arregladas)	Sustituciones por sitio después de 4 pasos (todas las mutaciones)	Sustituciones por sitio después de 4 pasos (mutaciones arregladas)
		I	II	III										
JEV	JEV I	DNS	DNS	DNS	BHK-21	BHK-21	1,32E+08	5,8	Sí	N/A	3,273	N/A	1,27E-03	7,29E-04
					SW13	SW13	1,52E+07	5,2	Sí	Positivo	0,409	N/A	7,29E-04	9,11E-05
					SW13*	SW13*	9,33E+06	2,8*	Sí	N/A	N/A	N/A	N/A	
	JEV III	I.C.	I.C.	I.C.	BHK-21	BHK-21	3,77E+07	6,1	Sí	N/A	3,286	1,143	1,54E-03	1,45E-03
					SW13	SW13	4,04E+06	4,8	Sí	Positivo	0,536	N/A	6,37E-04	-
	JEV I / JEV III quimérico	DNS	I.C.	I.C.	BHK-21	BHK-21	9,33E+07	6,7	Sí	N/A	0,404	1,571	1,36E-03	3,64E-04
					SW13	SW13	1,00E+07	6,8	Sí	N/A	1,19	1,589	9,10E-04	7,28E-04
	JEV III / JEV I quimérico	I.C.	DNS	DNS	BHK-21	BHK-21	6,58E+07	6,6	Sí	N/A	0,268	0,268	1,00E-03	2,73E-04
					SW13	SW13	3,06E+07	6,4	Sí	N/A	5,357	3,178	2,73E-04	6,38E-04
	WNV	Ouganda	I.C.	I.C.	I.C.	BHK-21	BHK-21	5,73E+07	5,3	Sí	N/A	0,268	N/A	4,55E-04
TBEV	Oshima	I.C.	I.C.	I.C.	BHK-21	BHK-21	3,28E+08	9,1	Sí	N/A	3,214	N/A	7,20E-04	9,00E-05
DENV-4	Dak HD 34 460	DNS	ARN vírico	DNS	SW13	VeroE6	6,59E+04	N/A	No	Positivo	0,436	0,535	8,45E-04	5,63E-04
YFV	BOL 88/1999	DNS	ARN vírico	DNS	SW13	HEK	1,40E+05	5,2	Sí	N/A	0,813	0,818	4,63E-04	4,63E-04
CHIKV	OPY1	I.C.	I.C.	I.C.	HEK-293	HEK-293	2,01E+07	7	Sí	N/A	2,24	N/A	4,12E-04	-
CV-B3	2679	I.C.	I.C.	I.C.	SW13	BGM	4,64E+07	7,4	Sí	N/A	N/A	N/A	2,70E-04	-
CV-B3*	2679*	No obtenido por PCR*			SW13*	BGM*	9,33E+07	7,4*	Sí	N/A	N/A	N/A	-	-

Tabla 3: Caracterización de los virus recuperados

Resumen de los diferentes virus producidos en este estudio: el nombre específico de la cepa, el origen del material inicial (DNS, síntesis *de novo*; I.C., Clon infeccioso; o ARN vírico) usado como el molde para la producción del primer (I), segundo (II) y tercer (III) fragmentos, la línea celular usada para la transfección y los pases, la cuantificación relativa de la cantidad de ARN vírico y títulos infecciosos en los sobrenadantes celulares en el cuarto pase mediante RT-PCR en tiempo real y ensayo TCID₅₀, la presencia o ausencia de efecto citopático (CPE), así como la investigación de antígenos víricos mediante ensayo de inmunofluorescencia directa (dIFA). Se obtuvieron secuencias completas del genoma vírico usando tecnología NGS.

dN y dS corresponden respectivamente al número de sustituciones no sinónimas por sitio no sinónimo y al número de sustituciones sinónimas por sitio sinónimo.

* Resultados obtenidos por transfección de seis fragmentos solapantes.

Resultados obtenidos transfectando directamente el clon infeccioso portador del plásmido CV-B3.

N/A y AU significan no disponible y unidad arbitraria, respectivamente.

En tercer lugar, los inventores demostraron la capacidad del método ISA para generar virus modificados genéticamente en días. Esto fue ejemplificado por la corrección basada en PCR de una mutación de cambio de marco (1915del) en el fragmento uno de un clon infeccioso JEV III defectuoso y la posterior recuperación del virus correspondiente (Métodos Suplementarios). También pudieron producir virus quiméricos intercambiando el primer fragmento de ADN (que codifica proteínas estructurales) de los JEV del genotipo I y III. A pesar de 11 emparejamientos erróneos en la región superpuesta de los dos primeros fragmentos, la transfección dio como resultado la producción de quimeras intergenotípicas JEV I/JEV III y JEV III/JEV I. El análisis de secuencias genómicas completas establecidas en el cuarto pase, usando NGS, mostró que la deriva genética (tasa de cambio de secuencia) era modesta (variando de 1,45E-03 a 9,00E-05 sustituciones por sitio cuando se consideran mutaciones fijas). La mayoría de mutaciones no sinónimas, la presencia de mutaciones compartidas entre las diferentes cepas de JEV (7/85) y la distribución no aleatoria de mutaciones (con una frecuencia superior al 10 %) a lo largo del genoma (con puntos calientes y regiones altamente conservadas) denotaron adaptación a las condiciones de cultivo celular.

La tasa de mutación varió de acuerdo con las células usadas para el aislamiento y, como se esperaba, fue más alta en los virus derivados de cepas de bajo pasaje que en los derivados de cepas adaptadas al cultivo. En conclusión, el método ISA es un procedimiento muy simple con el que se acelera la producción de virus ARN modificados genéticamente infecciosos en cuestión de días, con un control perfecto de las secuencias víricas y a partir de una diversidad de fuentes iniciales, incluyendo los clones infecciosos preexistentes, el ARN vírico o secuencias genómicas de ADN sintetizadas *de novo*. Esta técnica tiene el potencial futuro de generar el diseño de grandes experimentos de genética inversa para virus de ARN, en una escala que antes no se hubiera podido considerar. También tiene la capacidad, específicamente para modular las características de los virus recuperados de procedimientos experimentales. Adicionalmente, debido a que los fragmentos subgenómicos de ADN pueden obtenerse convenientemente mediante PCR, este método tiene el potencial de conservar la diversidad genética de poblaciones víricas 13 cuando se parte de ARN vírico. La PCR propensa a errores también puede usarse para crear heterogeneidad vírica artificial, por ejemplo, para facilitar la selección de virus adaptados bajo diversas condiciones de selección experimental y, por el contrario, pueden usarse polimerasas de alta fidelidad y moldes de amplificación clonal para controlar el grado de clonalidad de los virus producidos.

Finalmente, el método de la invención tiene el potencial de revolucionar la seguridad de los futuros intercambios de virus de ARN entre instituciones científicas, por el envío separado a temperatura ambiente de fragmentos subgenómicos de ADN sencillos, no infecciosos que, podrían combinarse y transfectarse después por el instituto receptor, permitiendo una recuperación rápida, sencilla y segura de la cepa vírica infecciosa.

5

Ejemplo 2: Atenuación de Chikungunya

MATERIALES Y MÉTODOS

10 *Células y anticuerpos*

Se cultivaron células de riñón de mono verde africano (Vero) a 37 °C con CO₂ al 5 % en un medio esencial mínimo (Invitrogen) con suero bovino fetal inactivado por calor al 7 % (FBS; Invitrogen) y Penicilina/Estreptomina al 1 % (PS; 5000 U/ml y 5000 µg/ml; Invitrogen). Se cultivaron células de riñón embrionario humano 293 (HEK293) a 37 °C con CO₂ al 5 % en medio de Eagle modificado de Dulbecco (Invitrogen) con FBS al 10 % y PS al 1 %. Se cultivaron células de *A. albopictus* C6/36 a 30 °C en medio L-15 (Invitrogen) con FBS inactivado por calor al 10 %, PS al 1 % y caldo de fosfato de triptosa al 5 % (29,5 g/l; Sigma-Aldrich).

15

20

25

Se usó un suero humano inmune específico de CHIKV para realizar el ensayo ELISA (véase a continuación). Para disminuir la concentración de moléculas inespecíficas que reaccionan con los compuestos celulares HEK293, se pusieron en contacto 40 µl de suero durante 16 horas con células HEK293 extraídas (células obtenidas de un cultivo en matraz de 150 cm², extraído usando acetona) en un volumen final de 400 µl (diluyentes: BSA al 1 %; KPL). Una proteína recombinante (fusión entre la región C-terminal de nsP2 y la región N-terminal de nsP3; Text S2), amablemente proporcionado por el laboratorio de AFMB (Architecture et Fonction des Macromolécules Biologiques, UMR 6098, Marsella, Francia), se usó para inmunizar dos conejos mediante métodos convencionales (protocolo de inmunización Rabbit Speedy de 28 días, Eurogentec). Los anticuerpos policlonales purificados (purificación por afinidad usando una matriz de sefarosa; Eurogentec) se usaron para realizar el análisis de transferencia Western.

30

Método de recodificación in silico

Se recodificaron tres regiones del genoma de CHIKV usando un programa informático que atribuía codones de nucleótidos aleatoriamente en función de su secuencia de aminoácidos correspondiente: por ejemplo, el aminoácido valina se reemplazó aleatoriamente por GTT, GTC, GTA o GTG. Para minimizar la influencia de codones raros en líneas celulares de primates, el número y la posición de tales codones raros en los genomas de primates (es decir, CGU, CGC, CGA, CGG, UCG, CCG, GCG, ACG) no se modificaron. Adicionalmente, se conservaron sitios de restricción únicos corrigiendo mutaciones sinónimas en algunos sitios. La ubicación de los casetes recodificados, en primer lugar, basándose en la disponibilidad de sitios de restricción únicos, se ajustó para evitar la superposición con estructuras secundarias de ARN conocidas. Finalmente, tres casetes de 1302, 1410 y 1500 bases y ubicados en las regiones nsP1, nsP4 y E2/E1, respectivamente, se diseñaron usando este método (las secuencias de los casetes se representan respectivamente en SEQ ID NO: 63, SEQ ID NO: 64 y SEQ ID NO: 65).

35

40

Construcción de clones infecciosos (IC) de CHIKV

Los presentes inventores modificaron un IC descrito previamente de la cepa LR2006 (registro de GenBank EU224268) reemplazando el origen de replicación y el promotor procariota por un origen pBR322 modificado y un promotor CMV (pCMV), respectivamente. Se usaron sitios de restricción únicos BamHI y XhoI para obtener un plásmido intermedio usando técnicas moleculares convencionales que contenían un nuevo origen de replicación (pBR322 modificado), el promotor procariota CMV (pCMV) y el genoma vírico parcial (desde la primera base hasta XhoI). El genoma vírico parcial (desde XhoI hasta el final), la cola poliA y la ribozima de la hepatitis D (HDR) seguida de una poliadenilación del virus de simio 40 (SV40) se sintetizaron (Genscript) y se introdujeron en la construcción intermedia usando sitios de restricción únicos XhoI y AvrII. Finalmente, se usaron sitios de restricción únicos BamHI, AgeI y XhoI para introducir mutaciones sinónimas en el genoma (se obtuvieron casetes mutados por fusión de productos de PCR). Se introdujo un total de ocho mutaciones sinónimas para generar los sitios de restricción requeridos o para eliminar los sitios de restricción indeseables. El clon infeccioso obtenido, que se consideró el tipo silvestre (TS), incorporó cuatro nuevos sitios de restricción únicos.

45

50

55

Todas las regiones recodificadas se sintetizaron (GenScript) y luego se insertaron en los circuitos integrados por digestión (BamHI/XmaI para ΦnsP1, AgeI/ApaI para ΦnsP4 y XhoI/AvrII para Φenv; NewEngland Biolabs), purificación en gel de productos de digestión (Qiagen), ligación (ADN ligasa T4; Invitrogen) y transformación en células STBL4 electrocompetentes (Invitrogen). Antes de su transfección, todos los clones infecciosos se purificaron (filtración de 0,22 µm) y se verificó su integridad mediante mapa de restricción y secuenciación completa usando un conjunto de pares de cebadores específicos.

60

Ensayos de RT-PCR en tiempo real

65

Se usó un fragmento de 179 nt ubicado en la región nsP2 (posición de nucleótidos 2631 a 2809) para detectar el ARN

genómico (cadena positiva) de todos los CHIKV (ensayo universal), recodificado o no. Se usó otro fragmento de 168 nt ubicado en la región nsP4 (posición de nucleótidos 6804 a 6971) para analizar los sobrenadantes celulares de los experimentos de competencia: dos conjuntos de cebadores y sondas permitieron detectar específicamente los virus recodificados en la región nsP4 o los virus sin modificación en la misma región.

5

Cinética de replicación

La aptitud replicativa de cada virus se determinó usando los resultados de los estudios de cinética de replicación, realizado por triplicado en células Vero, HEK293 o C6/36. Para comparar los siete virus de los siete IC (el virus TS y los 6 virus recodificados), se realizó un experimento con todos los virus. Se usaron reservas de virus o IC para infectar o transfectar células, respectivamente. Para la evaluación de la aptitud replicativa de los virus de pasaje, los inventores realizaron un experimento para cada virus (virus TS, Φ_{nsP4} y $\Phi_{\text{nsP1}} \Phi_{\text{nsP4}} \Phi_{\text{env}}$) con el primer pase en Vero y los pases 12, 25, 37 y 50 para cada régimen de pases (13 sobrenadantes probados por triplicado). Para la cinética de replicación de ciclo único, se usó una MOI estimada de 5 para infectar un matraz de cultivo de 75 cm² de células Vero, C6/36 o HEK293 confluentes. Las células se lavaron dos veces (HBSS) 30 minutos después de la infección y se añadieron 20 ml de medio. Se tomó una muestra de 1 ml de sobrenadante celular justo antes de los lavados y a las 2, 8, 14, 20 y 28 horas pi. Para la cinética de replicación con MOI estimada baja y la evaluación de la aptitud replicativa de los virus pasados, se usó una MOI estimada de 0,01 para infectar un matraz de cultivo de 25 cm² de células Vero o C6/36 confluentes. Las células se lavaron dos veces (HBSS) 2 horas después de la infección y se añadieron 8 ml de medio. Se tomó una muestra de 1 ml de sobrenadante celular después de los lavados (T0) y a las 24, 48 y 72 horas pi. Para la cinética de replicación usando clones de ADN infecciosos, se transfectó un matraz de cultivo de 75 cm² de células HEK293 subconfluentes con los IC usando Lipofectamine 2000 (Invitrogen) de acuerdo con las instrucciones del fabricante. Las células se lavaron dos veces (HBSS) 4 horas después de la transfección y se añadieron 20 ml de medio. Se tomó una muestra de 1 ml de sobrenadante celular después de los lavados (T0) y a las 16, 24 y 48 horas pi.

25

Todos los sobrenadantes celulares muestreados se clarificaron mediante centrifugación, se tomaron alícuotas y se almacenaron a -80 °C. Después se analizaron usando un ensayo TCID50 y un ensayo RT-PCR en tiempo real (no se realizó de forma sistemática, véanse las leyendas de las figuras). Los ácidos nucleicos se extrajeron de los sobrenadantes de células clarificados usando el EZ1 Virus Mini Kit v2 en el EZ1 Biorobot (ambos de Qiagen).

30

Experimentos de competencia de virus

El virus TS se cultivó en competencia con uno de los cuatro virus recodificados (Φ_{nsP4} , $\Phi_{\text{nsP1}} \Phi_{\text{nsP4}}$, $\Phi_{\text{nsP1}} \Phi_{\text{env}}$ o $\Phi_{\text{nsP4}} \Phi_{\text{env}}$) usando cinco proporciones diferentes de PFU (TS/virus recodificado 1/99, 20/80, 50/50, 80/20, 99/1). Se usó una MOI global estimada de 0,5 para la primera inoculación. Para cada experimento, un cultivo en matraz de 25 cm² de células confluentes se infectó durante 2 horas, se lavó (HBSS) y luego se incubó durante 48 h después de la adición de 7 ml de medio. Los virus de cada experimento se pasaron después nueve veces de la siguiente manera: un cultivo en matraz de 25 cm² de células confluentes se infectó durante 2 horas con el sobrenadante de cultivo purificado (centrifugación), se lavó (HBSS) y luego se incubó durante 48 h después de la adición de 7 ml de medio. En cada pasaje, la MOI estimada tuvo un cuello de botella en aproximadamente 0,5. Después de cada infección, los ácidos nucleicos se extrajeron del sobrenadante de células clarificado usando el EZ1 Virus Mini Kit v2 en el EZ1 Biorobot (ambos de Qiagen). Usando dos ensayos específicos de RT-PCR en tiempo real dirigidos a la región Φ_{nsP4} (véase anteriormente), se evaluó la cantidad de cada virus y se calculó la proporción de los dos valores (TS/recodificado).

45

Cuantificación de ARN intracelular y proteínas víricas

Se usó una MOI estimada global de 5 para infectar placas de 12 pocillos confluentes de células HEK293 con reserva de virus. Las células se lavaron una vez (HBSS) 30 minutos después de la infección y se añadieron 2 ml de medio. A 8 horas pi, se comprobó la ausencia de efecto citopático, se descartaron los sobrenadantes del cultivo y las células se lavaron una vez (HBSS). Todos los experimentos se realizaron por triplicado. Para análisis de transferencia Western y cuantificación de ARN vírico intracelular, el aislamiento total de ARN y proteínas se realizó usando el mismo pocillo con el kit Nucleospin ARN/proteína de acuerdo con las instrucciones del fabricante (Macherey-Nagel). Los extractos de proteínas se resolvieron en geles de poliacrilamida al 10 % que contenían SDS y se transfirieron a una membrana de PVDF. Se usaron pAb de conejo anti-Nsp1/2, mAb anti-actina C-2 (Santa Cruz Biotechnology) y el correspondiente anticuerpo secundario conjugado con HRP. Se revelaron bandas de proteína usando Immobilon (Millipore) seguido de exposición de transferencia a película radiográfica. Se realizó un ensayo de RT-PCR en tiempo real (véase anteriormente) para evaluar el ARN intracelular vírico (se usó actina de ARNm como normalizador para tener en cuenta las diferencias en el número de células y/o la calidad del ARN extraído como se describió anteriormente). Para la cuantificación de proteínas víricas mediante ELISA, las células se recolectaron mecánicamente usando un raspador de células, se resuspendieron en 800 μ l de PBS, se agitaron en vórtex y se rompieron sometiendo a ultrasonidos (30 segundos a 20 KHz, Misonix Sonicator XL). Se usó suero humano inmunitario específico de CHIKV pretratado para detectar proteínas víricas.

65

Paso experimental de virus in cellulo

El TS y dos virus recodificados (Φ ns_p4 y Φ ns_p1 Φ ns_p4 Φ env) se pasaron 50 veces siguiendo tres regímenes: pasajes seriados en células Vero o C6/36 y pasajes alternos entre Vero y C6/36. Para cada pasaje, se infectó un matraz de cultivo de 25 cm² de células confluentes durante 2 horas con el sobrenadante celular clarificado diluido, se lavó (HBSS) y se incubó durante 48 horas después de la adición de 7 ml de medio. Después se recogió el sobrenadante celular, se clarificó por centrifugación, se tomaron alícuotas y se almacenaron a -80 °C. Para cada pasaje, la MOI estimada tuvo un cuello de botella en aproximadamente 0,1. Para evitar la contaminación, los pasajes de virus se realizaron en tres fases: pasajes en serie de virus TS y Φ ns_p4, pasajes alternos de los mismos virus y pasajes del virus Φ ns_p1 Φ ns_p4 Φ env. Todos los virus pasados al mismo tiempo se manipularon secuencialmente y en diferentes cabinas de flujo laminar.

Ensayo en placa

Se infectaron monocapas de células Vero en placas de cultivo de 12 pocillos con 1 ml de reserva de virus (véase anteriormente). Después de dos horas, las células se lavaron (HBSS) y se añadieron 2 ml de agarosa al 0,9 % en medio de cultivo. Después de una incubación de 72 horas, las células se fijaron durante 4 horas con formaldehído al 10 % y se tiñeron durante 30 minutos con una solución de negro de naftaleno al 0,1 %.

Ensayo de dosis 50 infecciosa de cultivo de tejidos (TCID₅₀)

Para cada determinación, se inoculó un cultivo en placa de 96 pocillos de células Vero confluentes con 150 μ l/pocillo de diluciones seriadas de 10 veces de sobrenadantes de cultivo celular clarificados por centrifugación: cada fila incluía 6 pocillos de la dilución y dos controles negativos. Las placas se incubaron durante 7 días y se leyó la ausencia o presencia de CPE en cada pocillo. La determinación del DICT₅₀/ml se realizó mediante el método de Reed y Muench. Cuando el valor obtenido con una muestra fue menor que el umbral de detección del método (101,82 TCID₅₀/ml), los inventores realizaron otro ensayo con diluciones de dos veces, de 20 veces y de 200 veces (umbral de detección: 101,13 TCID₅₀/ml). Los valores inferiores a este umbral se consideraron iguales a 101,13 TCID₅₀/ml en las presentaciones gráficas y no se tuvieron en cuenta en los análisis estadísticos. Suponiendo que la recodificación y/o los pasajes experimentales podrían modificar significativamente la apariencia de CPE, los inventores usaron un ensayo de qRT-PCR (véase más adelante) como indicador sensible de la presencia de virus infecciosos. Este ensayo se realizó para cada virus (primer pasaje y, cuando estuviera disponible, 25 y 50 pases). Para todos los virus, los pocillos positivos para CPE fueron positivos en qRT-PCR con un ciclo de umbral inferior a 16 mientras que los que no produjeron CPE fueron negativos o positivos con un ciclo de umbral > 35, el valor esperado después de la dilución de los rendimientos iniciales de ARN.

Ensayo de hemaglutinación

Se usó una MOI estimada de 5 para infectar con stock de virus (ver arriba) un matraz de cultivo de 25 cm² de células confluentes Vero o C6/36. Las células se lavaron dos veces (HBSS) 30 minutos después de la infección y se añadieron 8 ml de medio sin FBS. Se tomaron muestras de 2 ml de sobrenadante celular a las 16 horas pi. Los sobrenadantes muestreados se clarificaron mediante centrifugación, se tomaron alícuotas y se almacenaron a -80 °C. Después se analizaron usando un ensayo TCID₅₀ (véase anteriormente), se realizó un ensayo de RT-PCR en tiempo real y un ensayo de titulación de hemaglutinación usando métodos convencionales: se prepararon diluciones en serie dobles de sobrenadante celular en microplacas con fondo en U en solución salina bovina al 0,4 %/solución salina de borato pH 9,0 (volumen final: 35 μ l por pocillo). Se añadieron treinta y cinco microlitros de glóbulos rojos de ganso pre-diluidos (1/150 usando los diluyentes de ajuste de pH 6,0 finales), la mezcla se homogeneizó, se incubó durante 45 min a temperatura ambiente y luego se leyó usando cuatro símbolos de puntuación: ++ para la hemaglutinación completa, + para hemaglutinación parcial, +/- para trazas de hemaglutinación y - para hemaglutinación negativa. El título de hemaglutinación fue el recíproco de la dilución más alta en la que se observó +.

RESULTADOS

Los inventores han evaluado el efecto sobre la aptitud replicativa y la citopatogenicidad de la recodificación a gran escala de CHIKV, un arbovirus patógeno reemergente del Viejo Mundo. La generación de virus atenuados mediante la codificación a gran escala representa una ruta emocionante y potencialmente importante para el desarrollo de vacunas y también para comprender la base de la evolución de la patogenicidad vírica. La recodificación dirigida al sitio, asociada a ninguna modificación de las secuencias de aminoácidos, alivia la probabilidad de novedosas propiedades fenotípicas, permite modular la aptitud alterando la longitud del intervalo de reemplazo de codones, pero adicionalmente aporta beneficios al desarrollo genérico de vacunas vivas atenuadas, incluyendo costes reducidos e inducción de una dosis única de inmunidad a largo plazo.

Un resultado clave fue la observación de que el método de recodificación aleatoria desvelado en el presente documento disminuyó la aptitud replicativa de CHIKV tanto en células de primates como de artrópodos. La disminución de la aptitud replicativa de CHIKV se correlacionó directamente con el grado de recodificación. Los inventores descubrieron que durante un ciclo de replicación en las células de mosquitos, la recodificación de codones redujo profundamente el título infeccioso del virus liberado mientras que el número de partículas víricas permaneció estable.

Esto implica que el proceso de maduración (es decir, la formación de ribonucleoproteínas y su inserción en las membranas plasmáticas que contienen HA) podría fallar cuando los virus se recodifican.

5 Por el contrario, en células de primates, esta disminución en la infectividad de las partículas víricas se asoció a la generación reducida de ARN vírico y proteínas probablemente debido a un complejo de replicación comprometido.

Estos resultados indican que las mutaciones sinónimas en los genomas víricos tienen importantes efectos de aptitud y no solo en el pequeño número de elementos que actúan en *cis* descritos anteriormente (Gerardin *et al.*, 2008).

10 De hecho, durante este experimento, se produjeron seis virus recodificados, de los cuales el virus más recodificado se modificó en tres regiones que codifican proteínas diferentes (en conjunto, se introdujeron 882 mutaciones sinónimas que abarcan 4.212 nt). En apoyo de estudios previos que demostraron que los virus de la gripe A recodificados y de poliovirus están atenuados, la observación de una reducción en la aptitud replicativa sugiere fuertemente que una proporción de mutaciones sinónimas no son neutrales en los virus de ARN. De hecho, es probable que algunas mutaciones sinónimas hayan sido seleccionadas positivamente durante el proceso de transferencia, reforzando la idea de que los sitios sinónimos son fundamentales para la aptitud vírica. En conclusión, es probable que las mutaciones sinónimas puedan ser neutrales, beneficiosas o perjudiciales como es el caso de mutaciones no sinónimas.

20 Los patrones evolutivos en sitios sinónimos podrían ser moldeados por procesos mutacionales en todo el genoma, tal como % de G+C, sesgo de uso de codones y frecuencia de dinucleótidos. Estas limitaciones globales, que teóricamente producen un subconjunto de genomas viables, se evaluaron por estudios previos de recodificación de codones en poliovirus, virus de la gripe A y virus bacteriano T7 que aplicó una modificación específica del sesgo de uso de codones, polarización del par de codones o frecuencias CpG/UpA.

25 Usando un método de recodificación aleatoria a gran escala, que solo modificó ligeramente estas propiedades globales, los inventores todavía observaron reducciones de la aptitud replicativa tanto en células de primates como de artrópodos. Estos resultados indican que las limitaciones locales también pueden proporcionar una presión de selección significativa en sitios sinónimos en virus de ARN, por ejemplo, interrumpiendo las estructuras secundarias de ARN. Dado que numerosas estructuras secundarias funcionales están presentes en las regiones codificantes de los virus de ARN y, por lo tanto, incluyen sitios sinónimos (con ejemplos notables en poliovirus, virus de encefalitis transmitida por garrapatas, alfavirus y VIH-1), es probable que estructuras similares sean comunes en CHIKV. Recientemente, se demostró que una estrategia de recodificación similar aplicada a las regiones sin cápside del poliovirus dio como resultado la identificación de dos nuevos elementos funcionales de ARN. El concepto de recodificación aleatoria a gran escala, como se describe aquí, también está respaldado por el informe del impacto negativo de mutaciones sinónimas únicas aleatorias (que no modificaron las características genéticas del genoma) sobre la capacidad de replicación vírica.

40 Finalmente, estos resultados indican que la reducción de la capacidad replicativa vírica está impulsada por una diversidad de factores.

Primero, la naturaleza del virus estudiado es un parámetro importante: los inventores encontraron que la introducción de hasta 882 mutaciones sinónimas aleatorias afectaba claramente la aptitud replicativa del CHIKV, mientras que dos estudios anteriores demostraron que los métodos de recodificación aleatoria comparables aplicados a la región del precursor de la cápside (P1) del poliovirus no afectaron significativamente la aptitud replicativa (se introdujeron 934 y 45 153 sustituciones sinónimos, respectivamente).

50 La ubicación de la región recodificada constituye el segundo factor de importancia: la recodificación en la región E2/E1 resultó en una mayor pérdida de aptitud que en otras regiones genómicas. El análisis de genomas completos de CHIKV de tipo salvaje reveló niveles naturalmente bajos de diversidad sinónima en esta región recodificada, lo que indica que esta región está sujeta a limitaciones evolutivas locales específicas que explican en parte el impacto significativo de la recodificación en esta región.

Es probable que el impacto medio de una mutación sea menos importante en la recodificación aleatoria que en enfoques específicos. Esto sugiere que la recodificación aleatoria a gran escala podría ser ventajosa en varios aspectos al diseñar futuras vacunas candidatas, en concreto:

- (i) la reversión al tipo silvestre debe ser intrínsecamente más difícil, dado el alto número de mutaciones producidas;
- (ii) dado que en los presentes experimentos la reducción de la aptitud replicativa disminuyó con el grado de recodificación, el método abre la puerta para ajustar con precisión la reducción de la aptitud mediante la modulación de la longitud de las regiones recodificadas y el número de mutaciones sinónimas introducidas;
- (iii) el uso de una combinación de varias regiones recodificadas ubicadas en todo el genoma vírico puede prevenir la reversión fenotípica completa debido a la recombinación entre virus TS y recodificados: la modificación de secuencia a gran escala puede hacer que la recombinación sea intrínsecamente más difícil y, en el caso de la recombinación, la parte del genoma que representa la cepa recodificada probablemente aún portaría algunas mutaciones asociadas a la reducción de la aptitud.

En consecuencia, estos virus recodificados son muy estables. Para corroborar, los inventores pasaron el tipo silvestre y dos CHIKV recodificados *in cellulo*. Durante el pasaje en serie de los virus recodificados, los inventores observaron que la respuesta a la recodificación de codones y la adaptación a las condiciones de cultivo se produjeron simultáneamente. Sin embargo, los altos niveles de evolución convergente observada entre el virus TS y los virus recodificados indican que la selección que surge de la recodificación del codón fue probablemente más débil que la de la adaptación a las condiciones de cultivo, y/o que las mutaciones beneficiosas para restaurar el coste de recodificación fueran menos probables. Por lo tanto, esta percepción indirecta de la dificultad de revertir los efectos de la recodificación destaca aún más la estabilidad de estos virus recodificados.

Estos experimentos también confirman que las mutaciones adquiridas en un hospedador pueden ser perjudiciales en un tipo de hospedador diferente (los pasajes en serie en células de primates aumentan la capacidad de replicación vírica en células de primates, mientras que los pasajes seriados en células de mosquitos disminuyeron la aptitud vírica en células de primates) y, con la excepción del virus más recodificado, que los pasajes alternativos (i) limitan seriamente la mejora de la aptitud replicativa y (ii) retrasan la aparición de las mutaciones.

En conclusión, estos experimentos demuestran que la recodificación de codones aleatorios disminuye significativamente la aptitud replicativa de CHIKV. Aunque todos estos resultados son importantes y alentadores, no pueden extenderse fácilmente a virus de ARN que producen infecciones crónicas. Por lo tanto, obviamente, se necesitan estudios en modelos animales para evaluar el potencial de estos métodos de atenuación de nueva generación para producir candidatos a vacunas. Sin embargo, este enfoque podría ayudar en el desarrollo de futuras vacunas de virus de ARN, incluyendo aquellas de los arbovirus. La introducción de una gran cantidad de mutaciones sinónimas ligeramente perjudiciales redujo la capacidad de replicación de CHIKV en órdenes de magnitud tanto en células de primates como de artrópodos. Esta estrategia dio como resultado una reversión limitada y una recuperación de la aptitud después de un subcultivo en serie intensivo de los virus, y es probable que reduzca el riesgo de reversión fenotípica completa si se produce la recombinación con el virus de tipo silvestre. Los presentes resultados alientan a que tales virus modificados tendrían dificultades para volver a su ciclo arboviral natural en el mundo real. Adicionalmente, la disminución de la aptitud replicativa se correlacionó con el grado de recodificación, una observación que puede resultar ventajosa en el desarrollo de estrategias futuras para modular la atenuación viral.

Ejemplo 3: Atenuación de otros virus de ARN

Se ha demostrado que la etapa de recodificación de codones a gran escala de la invención es un poderoso método de atenuación que tiene varias ventajas para el desarrollo de vacunas, incluyendo la posibilidad de obtener cepas de vacunas potenciales en un período muy corto tan pronto como se conozca la secuencia completa del patógeno diana y pueda producirse un genoma infeccioso. También tiene la posibilidad de modular con precisión el grado de pérdida de aptitud replicativa y generar vacunas seguras, vivas atenuadas que confieren protección a largo plazo, de una manera rentable.

Los inventores aplicaron el método de atenuación desvelado en el presente documento y ejemplificado en el ejemplo 2, a otros 2 arbovirus (ambos son flavivirus; virus de ARN de sentido positivo monocatenario envueltos): el virus de la encefalitis transmitida por garrapatas (TBEV) y el virus de la encefalitis japonesa (JEV).

A) TBEV

Siguiendo el método de recodificación de codones a gran escala aplicado previamente al virus Chikungunya (CHIKV), los inventores modificaron la región genómica NS5 (un casete de 1412 pb, como se representa en SEQ ID NO: 66) del TBEV (cepa Oshima 5-10), insertando 273 mutaciones silenciosas (recodificación de codones aleatorios).

La cepa TBEV Oshima 5-10, que se aisló en 1995 en Japón, pertenece al subtipo del Lejano Oriente y muestra una importante virulencia en ratones (provoca encefalitis como en humanos). Los virus de tipo silvestre (TS) y virus NS5_random_re-encoded se obtienen usando el método ISA y métodos clásicos (se obtuvieron clones infecciosos). Se midió la aptitud replicativa de los virus correspondientes *in cellulo* y era casi idéntica.

Tabla 4: Características genéticas del TBEV estudiado

Virus	Tamaño del casete (región genómica NS5)	Número de mutaciones	ENC del marco de lectura abierto completo	% G+C del marco de lectura abierto completo
TS	-	-	54,0	54,3
NS5 recodificado	1412 nt	273	55,5	53,8

El uso de codones se midió usando el número efectivo de codones ENC que da un valor que varía de 20 (solo un codón usado para cada aminoácido) a 61 (uso aleatorio de codones para cada aminoácido).

A continuación se usó un modelo *in vivo* para medir el fenotipo de atenuación de este TBEV recodificado. Después de la inoculación intraperitoneal (se usaron $2 \cdot 10^4$, $2 \cdot 10^5$ y $2 \cdot 10^6$ TCID₅₀ de virus), se controló la aparición de síntomas en los ratones y se pesaron todos los días durante 20 días.

5 Los resultados muestran un retraso en la pérdida de peso y la aparición de síntomas para los ratones infectados con el virus NS5_random_re-encoded en comparación con los infectados por el virus TS. Por otra parte, el número de ratones que muestran al menos un síntoma, pérdida de peso (≤ 94 %) y virus en el cerebro (detección de ARN vírico por RT-PCR en tiempo real) es significativamente mayor para los ratones infectados con TS que los ratones infectados por NS5_random_re-encoded. Se observaron altos niveles de anticuerpos IgG seroneutralizantes en ratones infectados con virus NS5_random_re-encoded a los 30 días después de la primera inoculación. Finalmente, los experimentos de desafío (los ratones fueron desafiados 30 días después de la primera inoculación) por el virus TS muestran que todos los ratones previamente infectados por virus recodificados estaban protegidos (basándose en la aparición de síntomas y pérdida de peso).

15 **Tabla 5: Características genéticas de los diferentes fragmentos TS, ligeramente recodificados o fuertemente recodificados.**

Virus	Fragmento I			Fragmento II			Fragmento III		
	Longitud	Mutación	% de G+C	Longitud	Mutación	% de G+C	Longitud	mutación	% de G+C
TS	3646	-	50,8	2854	-	52,3	3410	-	53,0
500	3646	225 (6,2 %)	49,7	2854	161 (5,6 %)	51,6	3410	199 (5,8 %)	52,0
1500	3646	672 (18,4 %)	49,1	2854	482 (16,9 %)	49,6	3410	563 (16,5 %)	50,3

Se indican el número del fragmento (primero, segundo o tercero), longitud, número de mutaciones sinónimas y % de G+C. 500 y 1500 significan un nivel bajo y alto de recodificación.

20 Usando el método de genética inversa ISA y combinaciones de estos fragmentos TS y recodificados, los inventores produjeron un gran número de virus recombinantes que albergan niveles graduales de recodificación en diferentes partes del genoma.

25 B) JEV

Los inventores han modificado la cepa JEV 'JEV_CNS769_Laos_2009' (Genotipo 1) usando el método de recodificación de codones aleatorios a gran escala.

30 Aquí se usa un enfoque diferente: los inventores recodificaron *in silico* casi todo el marco de lectura abierto (ORF) completo, desde el comienzo de PrM hasta el final de las regiones genómicas NS5, usando dos niveles diferentes de recodificación: un nivel alto y un nivel bajo de recodificación con la inserción de mutaciones sinónimas 585 o 1717 a lo largo de todo el marco de lectura abierto (**Figura 3**).

35 Para sus fines, los inventores usaron al menos un casete recodificado como se muestra en SEQ ID NO: 67; SEQ ID NO: 68; SEQ ID NO: 69; SEQ ID NO: 70; SEQ ID NO: 71; y SEQ ID NO:72.

40 La aptitud replicativa *in cellulo* de estos JEV recodificados se midió usando células humanas: Los resultados preliminares muestran una disminución de la aptitud replicativa según el nivel de recodificación, el tamaño de la región de recodificación y la posición genómica del fragmento o fragmentos recodificados (**Figura 4**).

Ejemplo 4: Generación *in vivo*

45 Se prepararon fragmentos solapantes que abarcaban todo el genoma de los virus de ARN y flanqueados respectivamente en 5' y 3' por el promotor de la ARN polimerasa dependiente de ADN y la señal de poliadenilación del terminador/ARN usando el método de la invención.

50 Estos fragmentos de ADN se inocularon directamente a animales vivos y se les permitió recuperar virus infecciosos de varias muestras de animales. Adicionalmente, la vigilancia clínica de los animales (aparición de síntomas y pérdida significativa de peso) permitió observar signos típicos de infección.

a) Experimento 1: Virus de la encefalitis transmitida por garrapatas (TBEV; *Flavivirus*)

55 Los inventores usaron una cepa de tipo silvestre del virus de la encefalitis transmitida por garrapatas (cepa Oshima 5.10 (número de registro de GenBank AB062063)). Aplicaron el método de la invención a fragmentos solapantes de ADN.

Se inocularon ratones hembra C57B1/6J de cinco semanas de edad con tres fragmentos solapantes de ADN.

El curso clínico de la infección vírica se controló siguiendo

- (i) las manifestaciones clínicas de la enfermedad (escalofríos, joroba, ojos sucios, hemi- o tetra-paresia, hemiplejía o tetraplejía); y
- (ii) el peso de los ratones exactamente como lo describe Fabritus L *et al.*, 2015, Attenuation of Tick-Borne Encephalitis Virus Using Large-Scale Random Codon Re-encoding. PLoS Pathog 11(3).

Se recogieron cerebros y bazo de ratones sacrificados 14 días después de la inoculación. Los cerebros y los bazo se molieron y se centrifugaron. El sobrenadante resultante se utilizó para evaluar la presencia de virus infecciosos.

La presencia de virus infecciosos se evaluó mediante métodos moleculares (RT-PCR en tiempo real) y métodos de cultivo celular clásico (aislamiento de virus infecciosos).

Usando una cantidad inicial de ADN que varía entre 2 y 5 µg y dos vías de inoculación diferentes (inyecciones intraperitoneales e intradérmicas), se detectaron virus infecciosos tanto en el cerebro como en el bazo. También se observaron manifestaciones clínicas (pérdida significativa de peso y síntomas) de las enfermedades.

b) Experimento 2: Inoculación intracerebral de ratones lactantes

Los inventores usaron cepas de tipo silvestre del virus de la encefalitis transmitida por garrapatas (cepa Oshima 5.10 (número de registro de GenBank AB062063)) y encefalitis japonesa (JEV_CNS769_Laos_2009 (número de registro de GenBank KC196115)). Usaron el método de la invención para generar los fragmentos solapantes de ADN.

Los fragmentos solapantes de ADN se usaron diluidos en PBS o se mezclaron con un reactivo de transfección.

Se inocularon ratones lactantes OF1 mediante inyección intracerebral de fragmentos solapantes de ADN. El curso clínico de la infección vírica se controló siguiendo la manifestación clínica de la enfermedad (escalofríos, letargo). Se recogieron cerebros de ratones sacrificados 6-12 días después de la inoculación. Los cerebros se molieron y se centrifugaron. El sobrenadante resultante se utilizó para evaluar la presencia de virus infecciosos.

La presencia de virus infecciosos se evaluó mediante métodos moleculares (RT-PCR en tiempo real) y métodos de cultivo celular clásico (aislamiento de virus infecciosos).

Usando 2 µg de ADN, se detectaron virus infecciosos en el cerebro para ambos virus (TBEV y JEV) y con o sin adición de reactivo de transfección. También se observaron manifestaciones clínicas de las enfermedades.

Conclusión

Los inventores han aprovechado así el poder de los métodos desvelados en el presente documento generando virus *in vivo*. Dicho método sería por lo tanto altamente eficaz para desarrollar una vacuna viva atenuada *in vivo*, es decir, directamente dentro del cuerpo un sujeto.

LISTADO DE SECUENCIAS

<110> Université d'Aix-Marseille

<120> Método para la generación rápida de un virus de ARN atenuado

<130> BCT150168 QT

<160> 72

<170> BiSSAP 1.3

<210> 1

<211> 781

<212> ADN

<213> Secuencia artificial

<220>

<223> Promotor

<400> 1

ES 2 873 083 T3

gaataagggc gacacggaaa tgtcacccaa ctgatcttca gcatcttcaa tattggccat 60
tagccatatt attcattggg tatatagcat aatcaatat tggctattgg ccattgcata 120
cgttgatctc atatcataat atgtacattt atattggctc atgtccaata tgaccgccat 180
gttggcattg attattgact agttattaat agtaatcaat tacgggggtca ttagttcata 240
gccatataat ggagttccgc gttacataac ttacggtaaa tggcccgctt ggctgaccgc 300
ccaacgacct ccgcccattg acgtcaataa tgacgtatgt tcccatagta acgccaatag 360
ggactttcca ttgacgtcaa tgggtggagt atttacggta aactgcccac ttggcagtac 420
atcaagtgta tcatatgcca agtccgcccc ctattgacgt caatgacggg aatggccccg 480
cctggcatta tgcccagtac atgaccttac gggactttcc tacttggcag tacatctacg 540
tattagtcac cgctattacc atggtgatgc ggttttggca gtacaccaat gggcgtggat 600
agcggtttga ctcacgggga tttccaagtc tccaccccat tgacgtcaat gggagtttgt 660
tttggcacca aatcaacggg gactttccaa aatgtcgtaa taaccccgcc ccgttgacgc 720
aatggggcgg taggcgtgta cgggtggagg tctatataag cagagctcgt ttagtgaacc 780
g 781

5 <210> 2
<211> 192
<212> ADN
<213> Secuencia artificial

10 <220>
<223> HDR - SV40pA

<400> 2

ggccggcatg gtcccagcct cctcgctggc gccggctggg caacattccg aggggaccgt 60
cccctcggta atggcgaatg ggactcgcga cagacatgat aagatacatt gatgagtttg 120
gacaaaccac aactagaatg cagtgaaaaa aatgctttat ttgtgaaatt aagcgtcggc 180
attgaccctg ag 192

15 <210> 3
<211> 23
<212> PRT
<213> Secuencia artificial

20 <220>
<223> cebador D

25 <400> 3

ES 2 873 083 T3

Cys Ala Cys Cys Cys Ala Ala Cys Thr Gly Ala Thr Cys Thr Thr Cys
 1 5 10 15
 Ala Gly Cys Ala Thr Cys Thr
 20

5 <210> 4
 <211> 25
 <212> ADN
 <213> Secuencia artificial

10 <220>
 <223> cebador I

<400> 4
 gaagaatgat tctgtaagtg tccag 25

15 <210> 5
 <211> 22
 <212> ADN
 <213> Secuencia artificial

20 <220>
 <223> cebador D

<400> 5
 cgttgccatg ccaatcttag cg 22

25 <210> 6
 <211> 22
 <212> ADN
 <213> Secuencia artificial

30 <220>
 <223> cebador I

35 <400> 6
 ggtgcttgcg tccttcacc aa 22

40 <210> 7
 <211> 25
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> cebador D

45 <400> 7
 caaatgagta tggaaatgctg gaaaa 25

50 <210> 8
 <211> 22
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> cebador I

55 <400> 8
 ctcagggtca atgccagcgc tt 22

60 <210> 9
 <211> 20
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> cebador D

5 <400> 9
 gccaccgga aggagctgac 20

<210> 10
 <211> 23
 10 <212> ADN
 <213> Secuencia artificial

<220>
 <223> cebador I

15 <400> 10
 cagagagcaa atccctatga cga 23

<210> 11
 <211> 22
 20 <212> ADN
 <213> Secuencia artificial

<220>
 <223> cebador D

25 <400> 11
 cgtcacatg ccagtcctag cg 22

<210> 12
 <211> 22
 30 <212> ADN
 <213> Secuencia artificial

<220>
 <223> cebador I

<400> 12
 35 gcttgcaat ccagtcagtc ct 22

<210> 13
 <211> 24
 40 <212> ADN
 <213> Secuencia artificial

<220>
 <223> cebador D

<400> 13
 45 caaacgagta cggaatgcta gaaa 24

<210> 14
 <211> 24
 50 <212> ADN
 <213> Secuencia artificial

<220>
 <223> cebador I

<400> 14
 55 ctcatgttg acagctatc atcg 24

<210> 15
 <211> 26
 60 <212> ADN
 <213> Secuencia artificial

<220>
 <223> cebador D

5 <400> 15
 tcaatattgg ccattagcca tattat 26

<210> 16
 <211> 24
 10 <212> ADN
 <213> Secuencia artificial

<220>
 <223> cebador I

15 <400> 16
 tggattgaac actcctgtag acgc 24

<210> 17
 <211> 23
 20 <212> ADN
 <213> Secuencia artificial

<220>
 <223> cebador D

25 <400> 17
 tggttggagt tgaagcctc atc 23

<210> 18
 <211> 19
 30 <212> ADN
 <213> Secuencia artificial

<220>
 <223> cebador I

<400> 18
 gaccatgccg tggccggcc 19

40 <210> 19
 <211> 24
 <212> ADN
 <213> Secuencia artificial

45 <220>
 <223> cebador D

<400> 19
 50 tggacaagac caagaatgac attg 24

<210> 20
 <211> 26
 <212> ADN
 55 <213> Secuencia artificial

<220>
 <223> cebador

<400> 20
 60 gttacaaata aagcaatagc atcaca 26

<210> 21
 <211> 24
 <212> ADN
 65 <213> Secuencia artificial

<220>
 <223> cebador

5 <400> 21
 cagggttatt gtctcatgag cgga 24

<210> 22
 <211> 22
 10 <212> ADN
 <213> Secuencia artificial

<220>
 <223> cebador

15 <400> 22
 gccacgcca ggaagagcat ga 22

<210> 23
 <211> 22
 20 <212> ADN
 <213> Secuencia artificial

<220>
 <223> cebador

25 <400> 23
 gggccctctg gaaatgggga ga 22

<210> 24
 <211> 24
 30 <212> ADN
 <213> Secuencia artificial

<220>
 <223> cebador

<400> 24
 caaccaggc ttgtaccat ctt 24

40 <210> 25
 <211> 22
 <212> ADN
 <213> Secuencia artificial

45 <220>
 <223> cebador

<400> 25
 50 gggtagagtc gtggaccttg ga 22

<210> 26
 <211> 28
 <212> ADN
 55 <213> Secuencia artificial

<220>
 <223> cebador

<400> 26
 60 cctaggaatt tcacaaataa agcatttt 28

<210> 27
 <211> 23
 <212> ADN
 65 <213> Secuencia artificial

	<220> <223> cebador	
5	<400> 27 cacccaactg atcttcagca tct	23
10	<210> 28 <211> 24 <212> ADN <213> Secuencia artificial	
15	<220> <223> cebador	
	<400> 28 gcatggaagt gtccttgag ttct	24
20	<210> 29 <211> 24 <212> ADN <213> Secuencia artificial	
25	<220> <223> cebador	
	<400> 29 gacttgcaac gatgctcttt tgca	24
30	<210> 30 <211> 22 <212> ADN <213> Secuencia artificial	
35	<220> <223> cebador	
40	<400> 30 gagagagcat cgtcacaatg cc	22
45	<210> 31 <211> 21 <212> ADN <213> Secuencia artificial	
50	<220> <223> cebador	
	<400> 31 gattccatcc agcaccgcac c	21
55	<210> 32 <211> 22 <212> ADN <213> Secuencia artificial	
60	<220> <223> cebador	
	<400> 32 ctcagggtca atgccagcgc tt	22
65	<210> 33 <211> 23 <212> ADN <213> Secuencia artificial	

<220>
 <223> cebador
 5 <400> 33
 gaataagggc gacacggaaa tgt 23
 <210> 34
 <211> 23
 10 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> cebador
 15 <400> 34
 tgaagacagc ttgtcctgca caa 23
 <210> 35
 <211> 24
 20 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> cebador
 25 <400> 35
 gatcatggct tggaggacca ttat 24
 <210> 36
 <211> 22
 30 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> cebador
 35 <400> 36
 gctactgcat agagcgtcca tg 22
 <210> 37
 <211> 25
 40 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> cebador
 45 <400> 37
 ttaccaggt aaaaacagaa accac 25
 <210> 38
 <211> 22
 50 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> cebador
 55 <400> 38
 ctcagggtca atgccagcgc tt 22
 <210> 39
 <211> 23
 60 <212> ADN
 <213> Secuencia artificial
 65

<220>
 <223> cebador

5 <400> 39
 cacccaactg atcttcagca tct 23

<210> 40
 <211> 23
 <212> ADN
 <213> Secuencia artificial

10

<220>
 <223> cebador

15

<400> 40
 catggaacca ttcctatgg act 23

<210> 41
 <211> 23
 <212> ADN
 <213> Secuencia artificial

20

<220>
 <223> cebador

25

<400> 41
 actggattgt gaaccaagga gtg 23

<210> 42
 <211> 25
 <212> ADN
 <213> Secuencia artificial

30

<220>
 <223> cebador

35

<400> 42
 gaagaatgat tctgtaagtg tccag 25

40

<210> 43
 <211> 22
 <212> ADN
 <213> Secuencia artificial

45

<220>
 <223> cebador

<400> 43
 cgttgccatg ccaatcttag cg 22

50

<210> 44
 <211> 22
 <212> ADN
 <213> Secuencia artificial

55

<220>
 <223> cebador

<400> 44
 aatataaccc cgagcggcga tg 22

60

<210> 45
 <211> 23
 <212> ADN
 <213> Secuencia artificial

65

<220>
 <223> cebador

5 <400> 45
 atgtcaccaa acaggtgcc caa 23

<210> 46
 <211> 22
 10 <212> ADN
 <213> Secuencia artificial

<220>
 <223> cebador

15 <400> 46
 ggtgcttgcg tcctccacc aa 22

<210> 47
 <211> 25
 20 <212> ADN
 <213> Secuencia artificial

<220>
 <223> cebador

25 <400> 47
 caaatgagta tggaatgctg gaaaa 25

<210> 48
 <211> 22
 30 <212> ADN
 <213> Secuencia artificial

<220>
 <223> cebador

<400> 48
 40 ggcctgtgct ccattgattc tg 22

<210> 49
 <211> 22
 <212> ADN
 <213> Secuencia artificial

45 <220>
 <223> cebador

<400> 49
 50 ggctgtgggc acattgtca cg 22

<210> 50
 <211> 22
 <212> ADN
 55 <213> Secuencia artificial

<220>
 <223> cebador

<400> 50
 60 ctcaggtca atgccagcgc tt 22

<210> 51
 <211> 23
 <212> ADN
 65 <213> Secuencia artificial

<220>
 <223> cebador

5 <400> 51
 cacccaactg atcttcagca tct 23

<210> 52
 <211> 21
 <212> ADN
 <213> Secuencia artificial

10

<220>
 <223> cebador

15

<400> 52
 ctgctcgggt gacctgtcct a21

<210> 53
 <211> 26
 <212> ADN
 <213> Secuencia artificial

20

<220>
 <223> cebador

25

<400> 53
 tgagatggtt ttctattca gcaact 26

30

<210> 54
 <211> 25
 <212> ADN
 <213> Secuencia artificial

35

<220>
 <223> cebador

<400> 54
 aacaatggtg tgacgaacag agtta 25

40

<210> 55
 <211> 22
 <212> ADN
 <213> Secuencia artificial

45

<220>
 <223> cebador

<400> 55
 ctccctgctg gacttgatag ag 22

50

<210> 56
 <211> 22
 <212> ADN
 <213> Secuencia artificial

55

<220>
 <223> cebador

<400> 56
 ctcagggtca atgccagcgc tt 22

60

<210> 57
 <211> 23
 <212> ADN
 <213> Secuencia artificial

65

<220>
 <223> cebador

5 <400> 57
 cacccaactg atcttcagca tct 23

<210> 58
 <211> 23
 <212> ADN
 <213> Secuencia artificial

10

<220>
 <223> cebador

15

<400> 58
 ccacacaaca tgcgtaccaa gca 23

<210> 59
 <211> 20
 <212> ADN
 <213> Secuencia artificial

20

<220>
 <223> cebador

25

<400> 59
 caggcgctgg cgctccgaca 20

30

<210> 60
 <211> 24
 <212> ADN
 <213> Secuencia artificial

35

<220>
 <223> cebador

<400> 60
 gtctatggtt atactctctg aaca 24

40

<210> 61
 <211> 22
 <212> ADN
 <213> Secuencia artificial

45

<220>
 <223> cebador

<400> 61
 gacaggagga cacaagtcag at 22

50

<210> 62
 <211> 22
 <212> ADN
 <213> Secuencia artificial

55

<220>
 <223> cebador

<400> 62
 ctcagggtca atgccagcgc tt 22

60

<210> 63
 <211> 1302
 <212> ADN
 <213> Secuencia artificial

65

ES 2 873 083 T3

<220>
 <223> Secuencias recodificadas

5 <400> 63

```

gatcctgatt ccaactattct agatataggg tctgcgccag caaggagaat gatgtcggac      60
agaaaatacc attgtgtttg tccgatgcgc agtgcggaag accctgagag actagcaaat      120
tatgcgagaa agctagcctc cgccgcaggg aaagtactgg ataggaatat ctctgggaaa      180
attggagacc tacaagcagt gatggcagtc cctgacacgg agacgcccac cttctgtctc      240
cacactgacg tatcttgacg gcaaagagct gatgtcgcga tctaccaaga tgtttatgca      300
gtgcatgcac ccacgtcgtt ataccaccaa gcgattaaag gtgtgcgagt agcgtactgg      360
gtagggttcg atactactcc gttcatgtat aatgccatgg cgggggcata tccaagctac      420
tcgacaaatt gggccgatga gcaggtgtta aaagccaaga atattggtct ttgcagcacc      480
gatttgacgg aaggtagacg aggaaaatta tctattatga ggggtaaaaa acttaaaccg      540
tgtgatcgtg ttctctttag tgtaggttca acgttgtacc cggagtcccg caagctcctt      600

aagagttggc atctgccctc ggtgttccac ttaaaaggta agctctcatt tacatgtcgc      660
tgtgacaccg tagtctcgtg cgaaggttat gttgtcaaaa ggataacgat gagccctggc      720
ctatacggaa aaacaaccgg atacgcggtc acccaccacg cagatggttt tctaattgtc      780
aagaccacag atacggtaga cggagagagg atgtcctttt cggtatgtac ttacgtgccg      840
gcgaccattt gcgatcagat gaccggcata ttggctacag aggtcacgcc ggaggatgct      900
cagaagttac ttgtgggttt aaatcaaagg atcgttgtaa acgggagaac gcaacggaac      960
acgaatacaa tgaaaaatta cttactacca gtggtcgctc aagcattctc caagtgggca     1020
aaagagtgtc ggaaagatat ggaggacgaa aaattattag gcgtcagggg aagaacactc     1080
acctgttgct gcctctgggc tttcaagaag caaaaaacac acacggttta caaaaggccc     1140
gacacacaat ccattcagaa agtacaggcc gaattcgatt ccttcgttgt gccgtcattg     1200
tggtcgtccg gtttaagcat ccccctcaga acaaggatta agtggcttct ctccaaggtt     1260
ccaaaaacag atctcatacc atactccggg gacgcccag ag                                     1302
    
```

10 <210> 64
 <211> 1410
 <212> ADN
 <213> Secuencia artificial

15 <220>
 <223> Secuencias recodificadas
 <400> 64

ES 2 873 083 T3

gccgtggccg cctgcaatga gttccttgca aggaactatc caacagtgtc cagttaccaa 60
 attactgacg aatatgatgc ttatctagat atggtagacg ggtcggaatc atgcttagat 120
 cgagcgacat ttaatccgtc taagttgagc tcttatccga aacaacatgc ataccacgcg 180
 cctagcataa gaagcgctgt accgagtcct ttccagaata ctctgcagaa cgtttttagcc 240
 gccgccacga aaaggaactg caatgttaca cagatgaggg aactcccaac attagactcc 300
 gctgtattta acgtcgaatg tttcaagaag tttgcatgca atcaagagta ttgggaagaa 360
 tttgccgcct cccctatcag aataaccaca gaaaatttag ccacttacgt cactaaatta 420
 aaagggcccta aggccgccgc gctgtttgct aaaactcata atttgcttcc cctgcaggaa 480
 gttcccatgg ataggtttac tgtcgatatg aaaagggatg taaaagtgac tccaggtacc 540
 aaacatactg aagaaaggcc aaaggtgcaa gtgatccaag cggcagagcc cttagcgaca 600
 gcctacttat gtggaattca tagagaattg gtgagaagat taaacgcagt cctcttgccct 660
 aacgtgcata ctcttttcga catgtccgct gaagacttcg atgccataat agcagctcat 720
 tttaaacccg gagatactgt gttggaaacg gatattgcat cattcgaaa aagccaagat 780
 gacagtttag cgttaacagc tttgatgttg ttggaagact taggggtcga tcactcctta 840
 ctagacctaa tagaagctgc cttcggggag atatcatctt gccatctacc gacaggtacg 900

 cgctttaaat ttggcgccat gatgaagagc ggcatgttct taacattatt cgtaaatacc 960
 ttgttaaaca ttaccatcgc aagtcgagtg ctagaggacc gttaactaa gagtgcgtgt 1020
 gcggcattta taggtgacga taacataatc catgggggtg tgagtgatga attaatggct 1080
 gccaggtgtg ctacctggat gaacatggag gtgaaaataa tcgatgcagt ggtctccctg 1140
 aaggcccat acttctgcgg gggttttatc ctccatgaca ctgttaccgg caccgcatgt 1200
 agagtagctg acccgtaaaa aaggttgttt aaactcggca agccgctagc ggctggggat 1260
 gaacaggacg aggataggag gcgagcgta gcagatgagg tgatcaggtg gcaacgaact 1320
 ggattgatcg acgagctaga gaaagcggtt tactcaagat acgaggtgca ggggattagt 1380
 gtggttgtga tgtcaatggc cactttcgcc 1410

<210> 65
 <211> 1500
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Secuencias recodificadas

<400> 65

5

10

ES 2 873 083 T3

gttacgtggg gaaataatga gccgtacaaa tattggccgc agttgtccac caacggtaca 60
gctcatggtc atccgcatga gattatactc tactattatg aactatatcc aacaatgact 120
gtcgttgtcg ttagtgtggc aacgttcata ctactttcga tggtaggtat ggctgcgggg 180
atgtgcatgt gcgcacgacg cagatgcata actccgtatg aattaactcc cggcgcaaca 240
gtacccttcc tactaagctt aatctgctgc atcaggacag caaaggcggc cacttaccag 300
gaggccgcga tttacctgtg gaatgaacaa caacctctgt tttggttga agcattaatc 360
ccgctagcag ctctcatagt tctgtgtaac tgtctgaggt tactcccctg ctgctgcaaa 420
acgttggcat tcctagcagt gatgtccgtg ggtgccca ca ctgtgtcagc gtatgaacac 480
gtgactgtca ttccgaatac ggtcggggtg ccgtataaaa ccttggtgaa tagaccggga 540
tactctccca tggttctgga aatggagctg ttgagcgtaa ctctggagcc tactctctcg 600
ctggattaca tcacgtgcga gtataagacc gtgataccga gcccgtagct gaagtgttgt 660
ggaactgcag agtgcaaaga taagaacttg cccgactact cttgtaaagt tttcacaggc 720
gtttaccctc tcatgtgggg tggggcatac tgtttttgtg acgcagaaaa tacgcagttg 780
tcagaagctc atgtagagaa gagcgagagc tgtaaaacag agttcgcttc cgcatacaga 840
gcccatactg catccgcttc cgcaaagcta cgcgttctat atcaaggcaa taatattact 900
gtgaccgcct acgccaacgg agatcatgca gttactgtta aagacgcaaa attcattgta 960
gggcctatga gcagcgcag gaccccctt gacaacaaga tagtggttta taaaggagac 1020

gtatacaata tggattatcc gccctttggt gctggcagac ctgggcaatt tggtgatadc 1080
cagtcacgca cccctgaaag taaggatggt tacgctaata cacagctcgt tctccagagg 1140
ccggcagtag gcacgggtaca tgtaccctac tcacaggccc cttctggttt taaatactgg 1200
ctgaaggaac gcggtgcgtc gttcaacat accgccccat tcggctgtca aattgccaca 1260
aaccgggtta gggcgggtcaa ctgcgctgta ggaaacatgc ctattagtat cgatatcccg 1320
gaggcggcat tcaccagggt ggtcgacgcg cctagtttaa cggatatgtc gtgcgaagtc 1380
cctgcttgta ctcacagtag tgatthtggg ggcgtagcaa tcataaagta cgctgcctcc 1440
aaaaaaggaa aatgtgcggt gcactcgatg acaaatgcag ttactattcg ggaagccgag 1500

<210> 66
<211> 1412
<212> ADN
<213> Secuencia artificial

5

<220>
<223> casete recodificado

10

<400> 66

ES 2 873 083 T3

ccgcggcgag ttaattaac ggagttgta agctcttacc atggccatgg aacgcgcggg 60
 aggatgtcgt gcgaatggcc atgaccgata ctaccgcctt tgggcagcag cgagtattca 120
 aagagaaggt agataccaag gccaggagc cccagccagg gacgaaggtg atcatgaggg 180
 ccgtcaatga ctggattcct gagcgacttg cccgaaagag taagcctcgg atgtgtagta 240
 gggaggaatt catagcgaag gtgaagagta acgcggctct tggggcctgg agcgatgagc 300
 aaaatagatg gtcatccgct aaagaggccg tcgaggaccc cgcattttgg caactgggtg 360
 acgaggaaag ggaaagacat ctggctggaa ggtgcgcaca ttgtgtctat aacatgatgg 420
 ggaaaagggg gaaaaagctt ggagagtttg gtgttgctaa ggggagtcgg gccatttggt 480
 acatgtggct gggcagccgc ttccttgagt tcgaagcact tggatttcta aacgaggacc 540
 actgggcttc caggggtcc agtggatcag gagttgaagg tatctcccta aattatttag 600
 gatggtacct aaagggttg agcactcttg agggcggact tttttacgca gatgacacag 660
 ccgatggga taccaagtc actaatgcag acttgaggga tgaagagcag ctccctacgtt 720
 acatggaggg tgaacacaag caactggcgg ctacaataat gcagaaggca taccacgcca 780
 aagtggtaaa agttgcccg ccctcccgag atggaggctg catcatggat gtcatcacta 840
 gaagagacca aagaggctct ggccaagtag tgacttatgc cctaaacacc ctaccaata 900
 ttaaagtaca actgatacga atgatggaag gcgagggtgt catcgaagca acggacgccc 960
 ataaccaag actgtttcga gtggaacgat ggctcagggg tcacggggag gaacgtcttg 1020
 ggagaatggt agtttccgga gatgactgtg tagtcagacc tgtcgtatgac aggttcagta 1080
 gagcgctata tttctgaac gatatggcca aaacaagaaa ggatgtaggc gagtgggaac 1140
 actcgggtggg tttctcgaat tgggaggagg ttcctttttg cagtcatcat tttcacgaat 1200
 tagtgatgaa agatgggcgc gccttaatag tgccttgccg agaccaagat gaattgggtg 1260
 gaagggcccg cgtctcccct gggtgccgct ggtcagttcg tgagacagcc tgtttgtaa 1320
 aggcataatg ccaaatgtgg cttttatcct attttcatcg gcgcgatctc cgaacgttag 1380
 gtttcgctat ctgttcggcg gtccccgtcg ac 1412

<210> 67
 <211> 3569
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> casete recodificado

<400> 67

5

10

ES 2 873 083 T3

ggaggggaatg aaagctcgat aatgtggcctt gcctcattgg caattgtaac agcttgtgca	60
ggagccatga aactatcaaa ctttcaagga aagttactga tgaccatcaa taacacggac	120
atagcggacg tcatcgttat ccccacctca aagggtgaaa acagatgctg ggtccgagca	180
atcgacgctg gatacatgtg tgaagacaca atcacgtatg aatgtccgaa gcttgccgtg	240
ggcaatgatc cggaagacgt cgactgctgg tgcgacaatc aggaagtcta cgtgcaatat	300
ggacgctgca cacggaccag gcattccaaa cgatccagaa gatccgtttc ggtccaaacg	360
catggggaaa gttcactagt taacaaaaaa gaggcttggc tggactcaac gaaggcaacg	420
cgatacctca tgaaaacgga aaactggatc ataaggaacc caggttatgc cttcttggcg	480
gcggcacttg gatggatgct tggctcaaac agtggacaac gtgtggtggt caccatcctc	540
ttattgttgg tcgctccggc ttacagcttc aactgtctgg gaatggggaa tcgggatttt	600
atagaaggag ccagtggagc cacttgggtg gatctggtgt tggaaggaga ttctgtctt	660
acaatcatgg caaacgacaa gccaacacta gatgtccgca tgatcaacat tgaagccagc	720
caacttgctg aagtcagaag ttactgctat cacgcatcag tcaccgacat ttcaacggta	780
gctcgatgcc ccacgactgg agaagcccac aacgagaaac gtgcagatag cagctacgtg	840
tgcaaacaag gctttactga tcgcggatgg ggtaatggat gtggactttt tgggaaagga	900
agcattgaca catgcgcaaa atttagctgt accaataagg ccattggaag aatgatccaa	960
ccagagaata tcaagtacga ggttggcata ttcgtgcacg gaaccaccac ctoggaaaac	1020
catggaaatt attcagcgca agtaggagcg tctcaagcag caaagtttac tgtaacccca	1080
aacgctcctt caataaccct caagttaggt gattatggag aggtcacact ggactgtgaa	1140
ccaaggagcg gattaaatac tgaagcgttc tatgtgatga cagtgggctc gaagtcattc	1200
ctagtccaca gggaatggtt ccatgatctt tctcttcctt ggacgagccc aagcagcacg	1260

ES 2 873 083 T3

gcatggagaa acagagaact cctcatggaa tttgaagagg cacatgccac aaaacaatct 1320
 gtcgtagccc ttgggtcaca ggaaggtggc ctccatcaag cgctggctgg ggccattgtg 1380
 gtggagtact cgagctcagt gaagttgaca agtggacact tgaaatgcag gctgaaaatg 1440
 gacaaaactgg ctttaaaggg cacgacttat ggcatgtgta cagaaaaatt ctcgttcgcg 1500
 aaaaatccag cggacacagg ccatggaaca gtggtcattg agctaacata ttctggaagt 1560
 gatggtccct gtaaaattcc gattgtctca gtggcgagtt taaacgacat gaccccagtg 1620
 ggtaggctgg taacagtaaa ccccttcgtc gccacatcta gctccaactc aaaggttctg 1680
 gttgagatgg aacctccctt cggagacagc tatatcgtgg tcggtagagg ggataagcag 1740
 attaaccatc actggcacia agctggaagc acgctgggca aagctttctc aactaccttg 1800
 aaaggggctc agagattagc agcgttaggt gacactgcct gggacttcgg ctccattgga 1860
 ggggtattta attctatagg gaaagctgtt caccaagtat tcggcgtgc tttcagaacg 1920
 ctctttggtg gaatgtcttg gatcactcaa ggactaatgg gggccctcct tttgtggatg 1980
 ggtgttaacg cacgagatcg gtcaatagct ctggcttttt tggccacggg aggtgtgctc 2040
 gtcttttttag cgaccaacgt gcatgccgac actggctgcg ccattgacat aactagaaaa 2100
 gagatgagat gcggaagtgg catctttgtg cacaacgacg tagaggcttg ggtggatagg 2160
 taaaaatata tgccagagac gcctagatcc ctagcgaaga tagttcacia ggcacatcaa 2220
 gagggagtgt gcggggtcag atccgttact agactcgaac atcagatgtg ggaatctggt 2280
 cgggacgaat taaatgtctt gctcaaagag aacgcggtcg atttgagcgt ggtcgtgaac 2340
 aaaccctggg ggagatatcg ctcagctccc aaacgcctat ccatgactca agaaaagttc 2400
 gagatgggct ggaaggcatg gggaaaaagc attctcttcg cccctgaatt ggccaactcc 2460
 acattcgtcg tggatggacc cgagacaaag gaatgccctg atgagcgcag agcttggaac 2520
 agcatgcaaa ttgaagattt cggattcggc ataacatcaa ctcgagtgtg gctgaaaatt 2580
 agagaggaga acaccgatga gtgtgatgga gcaatcatag ggacagcagt aaagggacat 2640
 gtggccgctc actccgactt gtcttactgg attgagagcc gtttgaatga cacctggaaa 2700
 ttggagaggg ctgttttcgg agaggtaaaa tcttgcaactt ggcccgaac acacactctt 2760
 tggggtgacg gagttgagga gagcgagctt atcatccac atactatagc tggaccgaga 2820
 agtaagcaca accggagaga aggtacaaa acacaaaacc agggaccctg ggatgagaac 2880
 ggcatcgtgc ttgactttga ttattgtcca ggaacaaagg tcacaatcac agaggactgt 2940
 ggcaaaaggg gtccctcaat cagaaccact actgacagtg gaaagctgat caccgattgg 3000
 tgctgtcgta gctgttctct accgccactc cggttccgga cagaaaatgg ttgctggtac 3060
 gggatggaaa tcaggcctgt taggcatgac gaaacaacac tagttaggtc acaggttgac 3120
 gctttcaacg gcgaaatgat tgaccattt cagctgggct tactggtgat gtttctcgca 3180

ES 2 873 083 T3

accaggagg tccttcgcaa gaggtggacg gccagattaa cgattccagc ggttctaggg 3240
gctctacttg tgctgatggt agggggcatc acttacactg acctggcaag atatgtcgtg 3300
ctagttgctg cggctttcgt ggaggccaac agtgggtggag atgttctgca tctcgtctg 3360
atagccgttt tcaaaatata gccagctttt ctggttatga atatgcttag cgcgaggtgg 3420
acgaaccaag aaaacgtggt tctggtcctg gggcggtt tttccaact cgcttcagtg 3480
gatttacaga tcggagtcca cggaatcctc aatgccctg ccatagcatg gatgatcgtt 3540
cgagcgatca catttcccac tacaagcac 3569

<210> 68
<211> 3569
<212> ADN
<213> Secuencia artificial

5

<220>
<223> Casete recodificado

10

<400> 68

ggtgggaatg aaagctcgat aatgtggctg gcctcaactcg ctattgtaac agcttgtgca 60
ggagcaatga aacttagcaa ttttcagga aaattattaa tgacaatcaa taacacggac 120
atagcggacg tgatcgttat acccacaagc aaggagaaa acagatgctg ggtccgagct 180
atcgatgtcg gatacatgtg tgaagacaca atcacgtatg agtgtccgaa gcttgccgtg 240
gggaatgatc cggaagatgt cgactgttg tgtgacaacc aggaagtta cgtccaatac 300
ggacgctgca cccggaccag gcacagcaa cgatccagaa ggagtgttc ggttcaaacg 360
catggggaga gttcactagt taataagaag gaagcctggc tggactcaac gaaggcaacg 420
cgatacctca tgaaaacgga aaactggata ataaggaacc caggatacgc cttcttggcg 480
gcggtccttg gatggatggt agggcacaac agtggacaac gtgtggtctt taccatcctc 540
ttattgcttg tagcaccggc ttacagcttc aactgcctgg ggatgggaaa tcgggatttt 600
attgaaggtg ccagtgggtc tacctgggtc gaccttgctc tcgaggggga ctctgtctt 660
accatcatgg ctaacgacaa gcctacactt gatgtccgca tgattaacat tgaagcctct 720
caactagctg aagtgaggag ctattgctat catgcatcag ttaccgatat cagtacggta 780
gcacgatgcc ccacgacagg cgaagcccac aatgagaaac gtgcagatag ctcttacgtg 840
tgcaagcaag gttttacaga ccgcgatgg ggtaatggat gcgggttgtt tggcaagggga 900
agcattgaca cctgcgcaaa attcagctgc accaacaagg ccattggaag aatgatccag 960
cccgagaata tcaagtacga ggtaggaata ttcgtgcatg gcaccactac atcggaaaat 1020
catggaaatt atagcgcgca agtaggcgag agtcaagcag caaaatttac cgtaaccccc 1080
aacgctcct caataacctt aaagttaggt gattatggtg aggtcactct agactgtgaa 1140

ES 2 873 083 T3

cccaggagcg gtttaaatac tgaggcgctc tatgtgatga cagttggctc gaagtctttc 1200
 ctagtgcaca gggaatgggt tcatgattta agccttccct ggacgagccc aagcagcacg 1260
 gcttgagaaa acagggagct cctcatggag tttgaagagg ctcacgccac taacaatct 1320
 gtcgtggccc tagggagtca ggaagggtga ctgcaccaag cgttggctgg ggccattggt 1380
 gttgaatact cgagcagtgt caaattgacc agtggacact tgaaatgcag gctgaaaatg 1440
 gataaactgg ctttaaaagg cacgacatat ggcatgtgca cagaaaaatt ctcgtttgcg 1500
 aagaaccccg cgacacaggg acatgggtact gtggttatcg agctaaccta ctcagggagt 1560
 gacggtcctt gtaaaattcc gattgtttct gtggcgagct tgaacgacat gacccagtg 1620
 ggtagactcg taacagtcaa cccattcgtc gccactagct caagtaactc caaagttctg 1680
 gtcgagatgg aacccccatt tggagatagc tatatagtgg tgggtagggg ggataagcaa 1740
 attaatcatc attggcacia agccggatcc acgttgggca aagctttctc aactaccttg 1800
 aagggggctc agagggttagc agcgctaggt gacactgctt gggactttgg ctccattgga 1860
 ggagtattha attctatagg gaaagcagtt caccaagtct ttggaggagc ttttaggacg 1920
 ctctttggtg gaatgtcatg gatcactcaa ggattgatgg gagcactcct gttgtggatg 1980
 ggcgttaacg cccgagatcg gtctatagct ttagcatttt tggccacggg ggcgtgctc 2040
 gtctttttg cgactaacgt gcacgctgac actggctgcg caatagacat aactaggaaa 2100
 gagatgagat gcggttccgg gatatttggt cataacgacg tggagcttg ggtggataga 2160
 tacaatatc tcccagaaac gcctagaagt ctggcgaaaa tagttcacia ggcacatcaa 2220
 gagggagttt gcgggggttag atctgttact aggctcgaac atcagatgtg ggaatctggt 2280
 cgggatgaat taaatgtctt gctaaaggag aacgcggtcg acttgagcgt ggtcgtgaac 2340
 aaacctgtgg gaagatatcg cagtgcctcc aaacgcctct ccatgacaca agaaaagttc 2400
 gagatgggct ggaaggcatg gggcaaatcc atcttgtttg caccagaatt ggccaacagt 2460
 accttcgttg tggatggacc cgagacaaaa gaatgccttg atgagcgcag agcttggaaat 2520
 tcaatgcaaa ttgaagattt tggattcggc ataacatcta ctcgagtttg gctgaaaatt 2580
 agggaagaga acaccgacga gtgtgacggg gcaattatag gcacagcagt aaaggacac 2640
 gtggccgttc actccgatct atcttactgg atcgaatctc gtttgaacga cacctggaag 2700
 ttggagaggg ctgttttcgg ggaggtaaaa agttgcactt ggccagaaac acataccctt 2760
 tggggcgacg gagtagagga aagtgagttg ataatcccc atactatagc tggccccgaga 2820
 agtaagcata accggagaga agggtaaaaa actcagaatc agggacatg ggacgagaac 2880
 ggaatcgtgt tagactttga ttattgtccc ggaactaagg ttacaattac agaggattgt 2940
 ggcaaaagag gtcctagcat cagaaccaca actgacagtg ggaagttaat aactgattgg 3000
 tgttgcgtt catgttccct gccgccactc cgttccgga cagaaaacgg ttgctggtac 3060

ES 2 873 083 T3

ggaatggaaa taaggcccgt tagacacgac gaaacaacac tagttaggtc acaggttgac 3120
gctttcaacg gcgagatgat tgaccattc caactgggct tactggtgat gttcctcgca 3180
actcaggagg tgttacgcaa gaggtggacg gctaggttaa cgattccagc ggttctaggc 3240
gcactgctcg tgctgatggt agggggaata acctacactg acctggctag atatgctgtg 3300
ctagttgctg cggccttcgt ggaggccaat agtggaggag atgttctgca tctcgccctg 3360
atcgagttt tcaaaataca gccagcattc ctcgttatga atatgctttc agcgaggtgg 3420
acgaaccaag aaaatgtggt tctagtcctg ggagcggctt tcttccaact cgcacccgtg 3480
gacttacaaa ttggtgttca cggcatcctc aatgctgctg caattgcttg gatgatcgta 3540
cgagcgatca cttttcccac tacaagcac 3569

<210> 69
<211> 2853
<212> ADN
<213> Secuencia artificial

5

<220>
<223> casete recodificado

10

<400> 69

ttgtcatagg aatttgttcc ctgctgcaag agaggagaaa gaccatggcg aagaagaaag 60
gagccgtgct cttgggctta gcgctcacat ctaccggatg gttctcgccc accactatcg 120
ctgccggact aatggtctgc aacccaaaca agaagagggg gtggccagcc accgaattcc 180
tttcagcggg tgggttgatg tttgccattg tgggagccct agccgagttg gacatcgaat 240
ctatgtcaat acccttcatg ctggcagggc tcatggcagt atcctacgtg gtatcaggaa 300
aggcaaccga catgtggctg gatcgggctg ctgatatcag ctgggagatg gaggcagcta 360
tcacaggaag tagccggagg ctagatgta agctggatga cgacggcgac tttcacttaa 420
ttgatgatcc cggagttcca tggaaagtct ggctcctccg catgtcttgt atcggattag 480
ccgctctgac accctgggct atcgtcccag ccgcttttggt ttattggctg accctgaaaa 540
caacaaaaag aggtggcgtg ttctgggaca cgccatctcc aaagccttgc ttaaaagggg 600
acaccaccac aggagtttac cgaattatgg ccagagggat tctcggcact taccaggccg 660
gtgtaggagt catgtacgag aacgttttcc atacattatg gcacacaact agaggcgag 720
ccatcatgag cggagaagga aagctaacgc catactgggg tagtgtgaag gaagaccgca 780
taagctatgg aggcccggtg aggttcgacc ggaatggaa tggaacagat gatgtgcaag 840
ttatagtggg ggaaccaggg aaacctgcag taaacattca gacaaaaccg ggtgtgtttc 900
gcaccctttt tggggagatt ggagccgtca gcctggacta cccacgggga acatccggct 960
ccccaatcct agattccaac ggagacatca taggcttgta cggcaatggg gttgagctcg 1020

ES 2 873 083 T3

gtgatgggtc gtatgtcagc gccattgtgc agggcgaccg tcaagaggaa ccagtcctcg 1080
atgcctatac tccaagcatg ctaaaaaaga gacagatgac tgtgctggat ctgcacccag 1140
gttcggggaa aaccaggaag atccttcccc aaataataaa ggatgctata cagcagcgtc 1200
tgagaacagc cgtgctggct ccaacacgag tcgtcgcagc tgagatggcg gaggccttga 1260
gaggtcttcc agtacgatat caaacctcag cagtgcagag ggagcatcag ggaaatgaaa 1320
ttgttgacgt aatgtgccat gccacactga ctcatagact aatgtctcca aacaggggtc 1380
ccaattacaa tttgttcgtt atggatgagg ctcaactcac tgatccagct agcatagccg 1440
ctcgggggta tatagcgacc aaggtggaac tgggggaggc agcagccatt tttatgacgg 1500
cgaccccgcc cgggaccact gaccctttc ccgattcaaa tgctcccatt cacgacctgc 1560
aggatgagat ccagacaga gcatggagta gtggttacga atggatcacg gattacgcgg 1620
gaaaaactgt atggttcgtg gcaagtgtta aaatgggaaa cgagatcgcc atgtgcctcc 1680
aaagagcggg aaaaaaggtc atccaattga accgtaagtc atatgacaca gagtacccaa 1740
aatgtaagaa tggagactgg gactttgtga tcaccactga catctctgag atgggggcca 1800
attttggtgc gagcaggggt attgactgca gaaagtcctg gaaacctacc atcctagagg 1860
agggagaggg tagagtcatc ctcgaaacc catccccat aaccagtgtc agtgcagccc 1920
agcggagagg aagagtggga aggaatcca accaagttgg cgatgagtac cattatggag 1980
gggccaccag tgaagatgat agcaacctg ccactggac agaggcaaaa attatgctag 2040
ataacataca catgcctaata ggcttagtgg ctcaactgta cgggcctgaa agggaaaagg 2100
ctttcacaat ggatggagaa taccgttga gaggtgagga gaagaagaac ttcttgagc 2160
tgcttagaac ggctgactta ccagtatggt tggcctacaa agtggcgtcc aatggcattc 2220
agtacacaga cagaaaatgg tgctttgatg gaccacgcac aaacgtata ctggaagata 2280
aactgaggt ggagatagtc acccgaatgg gtgagagaaa gatcctcaag ccgagatggc 2340
tcgatgcgag ggtttatgca gaccaccaag ctctcaagtg gttcaaagat tttgcagcgg 2400
gcaagagatc tgccgtctcc ttcatagagg tactcggtcg catgccagag catttcatgg 2460
gaaagacacg ggaggcctta gacacaatgt acctcgtggc aacagccgag aaaggcgga 2520
aggctcaccg catggctctt gaagaactgc ccgacgcatt ggagaccatc aactcatcg 2580
ttgccatcac tgtgatgaca ggaggathtt tcctgctcat gatgcaaaga aagggtattg 2640
gaaaaatggg ctaggggct ctagtgttga cgctggccac ctttttccta tgggcggctg 2700
aagtccttgg aactaaaatc gcgggcactc tactggttgc ctttttctg atggtggtcc 2760
tcatcccgga accagaaaaa cagaggtcac agacagacaa tcagttggca gtatttctta 2820
tctgcgtcct gactgtggtc ggagtgtggc cag 2853

<210> 70
<211> 2853
<212> ADN
<213> Secuencia artificial

ES 2 873 083 T3

<220>

<223> casete recodificado

5 <400> 70

ttgtcatagg aatttgttcc ctattgcaag aaaggagaaa aacctatggcg aagaaaaagg	60
gtgcagtgtt gcttggatta gcgctcacct ctactggctg gttctcgccc actaccatag	120
ctgccggtct aatggtatgc aatcctaaca aaaagagggg ttggcctgct accgaattcc	180
tgtctgcggt cgggttgatg ttcgccatag tgggcggcct agcagaactc gacatcgaat	240
caatgtctat tccctttatg ttggcagggc tcatggcagt aagctacgtc gtatcaggaa	300
aagctaccga catgtggctg gatcgggctg ctgatatatc ttgggagatg gaggcagcta	360
tcacaggaag tagccggaga ttagatgtaa aactggatga tgacggcgac ttccacttaa	420
tagacgatcc cggagtcca tggaaagtgt ggctcctccg catgagttgc ataggactgg	480
ccgctctgac tccttgggca atcgtcccag cagcctttgg ttactggctg accttaaaaa	540
ccacaaaaag aggtggcgtg ttctgggata cgccatctcc caagccctgt ttaaaaggag	600
ataccaccac aggggtctac cgaattatgg ctagaggaat tcttggcact taccaggcag	660
gtgtaggagt aatgtacgaa aacgttttcc ataccttatg gcacactacc agaggcgcag	720
ccattatgag tggagaagga aagctcacgc cctactgggg tagtgtaaaa gaagaccgca	780
ttagctatgg aggtccgtgg aggtttgacc ggaaatggaa cggtagacagat gatgtacaag	840
ttatagtggg ggaacccggg aaacctgcag taaacattca aactaaaccg ggtgtcttcc	900
gcactccatt tggggaaatt ggagccgtct ctctagatta tcctcggggg acaagcggct	960
ccccaatctt ggattccaac ggagatatta taggcttata cggtaatgga gtggagttgg	1020
gtgatgggtc gtatgtcagc gccatagtgc aaggcgatcg tcaagaggag ccagtcctctg	1080
atgcttatac tccaagcatg ctaaagaaga gacagatgac tgtgctggat cttaccctg	1140
gttcggggaa aactaggaag atacttcctc aaataataaa ggatgctata caacagcgtc	1200
tgagaacagc agtcctggct ccaacacgag tcgtcgcagc tgagatggcg gaggccctga	1260
gaggtcttcc cgtacgatac cagacctctg cagtgcagag agaacatcag ggcaatgaaa	1320
ttgttgacgt aatgtgccac gctacactta ctcataggtt aatgtctcca aataggttcc	1380
ccaattacaa tttgttcggt atggacgagg ctcaactttac cgatcccgtc agcatagccg	1440
ctcgggggta catagcgacc aaggttgagc tcggagagggc agctgccatt tttatgacgg	1500
cgactccgcc aggaacaact gacccttcc ccgatagtaa tgctcccatt cacgacctcc	1560
aggatgagat cccagacaga gcatggagta gtggttacga atggatcacg gattacgcgg	1620

ES 2 873 083 T3

gaaaaactgt atggttcgtg gcttcagtta aaatgggaaa cgagatcgcc atgtgtctcc 1680
 aaagagcggg aaaaaaagtc atccaattga accgtaagtc ctacgatact gagtacccca 1740
 agtgtaagaa tggggactgg gactttgtga taacaacaga cataagtgag atgggcgcca 1800
 acttcggcgc gagtaggggtt attgattgta gaaagtcogt gaaacctacc atactggagg 1860
 aaggcgaggg cagagtcatt ctcggaacc cttccccaat aactagtgct agtgcagccc 1920
 agcggagggg aagagtggga agaaatccca accaagttgg cgacgagtac cactatggag 1980
 gggccactag tgaggatgat tcaaatcttg cccactggac tgaggctaaa attatgctgg 2040
 ataatataca tatgcctaac ggccttggtg cacaactgta cgggcctgaa agggaaaagg 2100
 cctttacaat ggacggagaa taccgtttga ggggtgagga gaaaaagaac ttccttgagc 2160
 tgctgagaac ggctgactta cctgtttggg tggcctacaa ggtggcgctc aatggcatcc 2220
 agtacacaga tagaaagtgg tgttttgatg gaccacgcac taacgctatc ctggaagata 2280
 acaccgaggt agagatcgtg acccgaatgg gtgagaggaa gatcttaaag ccgagatggt 2340
 tagacgcgag ggtttatgca gatcaccaag ccctaaaatg gttcaaagac ttcgcagcgg 2400
 ggaagagatc tgccgtttcc ttcatagaag tattgggtcg catgccagag cacttcatgg 2460
 gaaagacacg ggaggctttg gacacatgt atctcgtagc aacagccgag aaaggcggaa 2520
 aggctcatcg catggcactt gaagagctgc ctgatgcaact ggaaactatc aactaatcg 2580
 ttgcaatcac agtgatgaca ggggggtttt tccttctgat gatgcaaaga aagggcattg 2640
 gaaaaatggg ccttggagct ctagtgttga cgttagcaac ttttttcta tgggcggtg 2700
 aagtccctgg aactaaaatc gcgggaactc tcctggttgc ccttctactg atggtcgtcc 2760
 tcataccgga accagaaaaa cagagatctc agacagacaa tcagttggct gtatttttaa 2820
 tctgctctt aactgttggt ggagtcgtcg cag 2853

<210> 71
 <211> 3409
 <212> ADN
 <213> Secuencia artificial

5

<220>
 <223> casete recodificado

10

<400> 71

aggtctgacc ggattgccta gcatggctct ggacttgccg ccagccacag cttgggctct 60
 gtatgggggt agcacagttg tgtaacccc cctcctgaag cacctaata cctcagaata 120
 cgttaccaca tcgtagcat caataagttc ccaagcgggt tcgctgtttg ttttgccgct 180
 cggcgtgcct tttactgact tggatctaac cgttggactt gtctttctcg gctgctgggg 240
 ccaaatacacc ctcaccacgt tcctaacagc tatggtacta gtaaccctcc actatggata 300
 catgctgcct ggttggaag cagaggcact aagagctgct cagagaagaa cagcggctgg 360

ES 2 873 083 T3

cataatgaag aatgcagttg tggacggaat ggtcgcaaca gacgtgcccg agcttgaaag 420
 aactactccc ttgatgcaaa agaaagtcgg gcaagtgctc ttgatagggg tgagcgtggc 480
 ggcgtttttg gtgaacccaa atgtcaccac cgtgagagag gcaggtgtat tgggtgacggc 540
 tgccacactc accttgtggg ataatggagc ctctgccgtc tggaacagca ccaccgccac 600
 ggggctttgc catgtcatgc gaggcagcta tctagccggc ggatccatcg cctggactct 660
 cattaanaac gctgataagc cctccttaa aagggggagg cctggtggca ggacgctagg 720
 ggagcagtgg aaggagaaac ttaacgctat gagcagggat gagttcttca aatacagaag 780
 agaggcaata attgaggtgg accgcactga agcagcagg gctaggcgcg agaacaacat 840
 agtgggagga catccagtct cgcgagggtc cgcaaagctc cgctggctcg tggaaaag 900
 atttgtcagt ccaataggaa aggtcataga tctgggggtc gggcgcgag gctggagcta 960
 ctacgcagca actctgaaga aagttcaaga agtcaaagga tacacgaaag gtggagcggg 1020
 acacgaagaa ccgatgctca tgcaaagtta cggttggaac ctggtctcgt taaagagtgg 1080
 ggtggacgta ttctacaaac cttcggagcc ttccgacacc ctgttctgtg acataggaga 1140
 atctagccca agtccagagg tggaggagca acgcacgctg cgcgtcctag aatgacatc 1200
 cgactggcta catcggggcc ctagagagtt ttgcataaaa gttctctgcc cttacatgcc 1260
 taaagttata gagaaaatgg aagtgttaca acgtcgtttc ggtggcggcc tgggtgcgcct 1320
 tccactttct cgaaactcta accatgagat gtattgggtg agtggagctg ccggcaatgt 1380
 tgtgcatgcg gtcaacatga ccagccaagt gttactagga cgaatggatc gcacagtgtg 1440
 gagaggtcca aagtatgaag aggatgtcaa tctgggcagc gggacgagag ctgtggggaa 1500
 gggggaggtc catagcaacc aggaaaaaat taggaagaga atccagaaac tcagagaaga 1560
 attcgcaaca acctggcaca aggaccctga gcaccatac cgaacctgga cctatcacgg 1620
 aagctacgaa gtgaaggcta ctggcagcgc aagctctcta gtcaacgggg tggtaaaact 1680
 catgtctaaa ccctgggatg ctatcgcaaa tgtcaccaca atggccatga cagacacaac 1740
 cccctttggc cagcagaggg tcttcaagga gaaggttgac acgaaggctc cagagccacc 1800
 agcaggagtt aaagaagtgc taaacgagac caccaactgg ctgtgggccc atttgtcacg 1860
 ggagaaacga cctcgcttgt gcactaaaga agaattcata aagaaagtga attctaacgc 1920
 agcttttagga gcagtgttcg ccgaacagaa tcaatggagc acggcgcggg aagctgtggg 1980
 tgaccccctg ttttgggaga tggatcaatga agagagagag aatcatttgc gaggcgagtg 2040
 ccatacgtgc atctacaaca tgatgggaaa aagagagaaa aaaccggag agttcgggaa 2100
 ggccaaaggg agtagggcta tttggttcat gtggctcggg gctcgggtacc tagagttcga 2160
 agccctagga tttttaaag aagaccattg gctttctcga gagaattcag gaggcggggg 2220

ES 2 873 083 T3

agaaggttca ggtgtgcaaa agttgggtta cattctccgg gacatagccg ggaagcaagg 2280
 cggtaaaatg tatgctgatg acacagccgg gtgggacacc agaatcacca gaaccgactt 2340
 agaaaatgaa gccaaagtgc ttgagctttt ggatgggtgaa caccgcatgc tcgcccgagc 2400
 cataattgaa ctaacgtaca ggcacaaaagt ggtaaagtt atgagacctg ccgcaggagg 2460
 aaagacagtg atggacgtga tatcccgaga agaccaaagg gggagtgggc aggtggtgac 2520
 ctacgctctc aacacattca cgaacattgc cgtccaactg gtccgcttga tggaggcaga 2580
 gggggtcatt ggacccaac acttgaaca gctgccaggg aaaaacaaa tagcagtcag 2640
 gacatggctc tttgagaatg gagaggagag agtgactaga atggcgatca gtggagacga 2700
 ctgtgtcgtt aagccgctgg atgacagatt cgccacggct ctccatttcc tcaatgcaat 2760
 gtcgaaagt t agaaaggata tccaggaatg gaagccttcg catggttggc acgactggca 2820
 gcaggtcccc ttttgctcca atcatttcca ggaaattgtg atgaaagacg gaaggagcat 2880
 agtcgtgccg tgcagagggc aggatgagct gattggcagg gcgcgcatct ccccaggtgc 2940
 tggatggaat gtgaaggaca cagcttgtct ggccaaagcg tatgcacaga tgtggctgct 3000
 cttgtacttc catcggaggg acctacgct catggcaaat gcaatttgca gtgcagttcc 3060
 agtggactgg gtgcccaccg gcagaacatc ctggtcaata catagcaaag gagagtggat 3120
 gaccactgaa gacatgctgc aagtatgga cagggatgg attgaggaga atgagtggat 3180
 gatggataaa accccaatca catcctggac agacgttccg tacgtaggaa agcgcgagga 3240
 catctggtgt ggcagtctca tcggaacgcg aagcagggca acatgggctg agaacatcca 3300
 cgcggctata aaccaagtga gggccatcat tggaaaagag aattatgttg attacatgac 3360
 ttccctcaaa agatatgagg atgtattgat ccaggaggac agggtcatt 3409

<210> 72

<211> 3409

5 <212> ADN

<213> Secuencia artificial

<220>

<223> casete recodificado

10

<400> 72

tggcttaacc ggccttcaa gcatggctct agacttgcgc ccagcaaccg cttgggctt 60
 gtacgggggt agcactgtcg tactcactcc cctcttaaag cacctaata cctcagaata 120
 cgttactact tcgtagcaa gcataagttc ccaggcgggc tcgctgttcg tactcccgcg 180
 cggcgtgccc ttcactgact tggatctaac tgtcggacta gtctttctcg gatgctgggg 240
 gcaaataacc ttgacaacgt tcctaacagc catggtactc gtaaccctcc attatgggta 300
 tatgctgcca ggttgcaag cagaggctct aagagcagcc cagagaagaa cagcggctgg 360
 tattatgaag aatgcagttg tggatggaat ggtcgcaca gacgtgccg agcttgaaag 420

ES 2 873 083 T3

aactacacca ttgatgcaaa aaaaggtggg gcaagtgcta ttgataggtg tgagcgtcgc 480
ggcgttcttg gtgaatccaa atgtgaccac cgtcagagaa gctgggggat tggttacggc 540
tgccactctt accctctggg ataatggggc atctgctgtc tggaacagca ccaccgccac 600
ggggctttgt catgtcatgc gaggctcata tttagccggc ggaagcatcg cctggactct 660
cattaaaaat gccgataaac catcattaa gagggggagg cccggtggga gaacgttagg 720
ggagcagtgg aaggagaaac ttaacgctat gtcaaggac gagttcttca aatatagaag 780
agaagcaatc attgaagtgg accgactga agcacgcaga gctaggcgcg aaaacaatat 840
agttggcggg catccagtgt cgcgagggtc cgcaaagctc cgctggctag ttgaaaaggg 900
ttttgttagt ccaataggga aggtgatcga tttgggttgc gggcgcgag gttggagcta 960
ctacgcagca actctgaaga aggttcaaga ggtaaaggga tacacgaagg gtggagcggg 1020
gcacgaggag ccgatgctca tgcaatctta tggatggaac ctagtatcgc ttaaagtgg 1080
ggtagacgta ttctataagc cttcggagcc ttccgacacc ctgttttgtg acatcggaga 1140
atctagccct agtcagagg ttgaggagca acgcacgta cgcgtcttag aaatgacatc 1200
cgactggcta catcggggcc caagagaatt ttgcatcaa gttctgtgtc catacatgcc 1260
taaggttatt gagaaaatgg aagtgttaca acgtcgtttc ggtgggggcc tcgtgcgcct 1320
tccactttcc cgaaatagta accacgagat gtattgggtg agtggcgctg ccggcaatgt 1380
tgtgcatgcg gtcaatatga ccagccaggt attactgggt cgaatggatc gcacagtctg 1440
gagagtcctt aaatatgaag aggatgtaa tctgggttct gggacgaggg ctgtagggaa 1500
gggggaggtc catagcaacc aggaaaaaat taggaagagg atccagaaac tcagagaaga 1560
attcgcaaca acatggcaca aggaccctga gcaccctat cgaacctgga cctatcacgg 1620
ctcatatgaa gtaaaggcta ctggaagcgc atccagccta gtaaaccggg tagtaaaact 1680
aatgtctaaa ccctgggatg ctattgcaaa cgtgaccacc atggcaatga cagacacaac 1740
ccctttcggc cagcagagag tattcaagga gaaggtcgac acgaaggctc ctgagccccc 1800
agcagggggt aaagaagtac taaatgaaac caccaactgg ctgtggggcc atctcagccg 1860
ggagaaacga cctcgcctgt gcactaaaga agagttcatt aagaaagtga attctaacgc 1920
tgcattagga gccgtattcg ccgaacaaaa tcagtggagc acggcgcggg aagccgttgg 1980
tgaccccctg ttttgggaga tggtaaacga ggagagagag aatcatttgc gaggcgaatg 2040
tcatacgtgc atttacaaca tgatgggcaa gagagaaaa aaacccggtg aattcgggaa 2100
agccaaaggt agtagggcaa tatggttcat gtggctcggc gctcggtagc tagaatttga 2160
agctttgggc tttttaaatg aagaccactg gctttctcga gaaaattctg gggcggtgt 2220
agaaggatca ggtgtgcaaa aattgggtta tatacttcgg gatatagccg ggaacaggg 2280

ES 2 873 083 T3

cggaaaaatg	tatgctgatg	acacagcagg	atgggacacc	agaatcacca	gaaccgactt	2340
agaaaacgaa	gccaaagtgc	ttgagctcct	tgatggtgaa	caccgcatgc	tcgcccgagc	2400
tatcattgaa	ctaactgaca	ggcataaagt	ggttaaagtt	atgagacccg	ccgcaggcgg	2460
aaagacagtg	atggatgtaa	tatcccgaga	ggaccaaagg	gggagtgggc	aggtggtgac	2520
atacgccctt	aacactttca	cgaacattgc	cgttcaactg	gtacgcctta	tggaggcaga	2580
aggtggtata	ggcccccaac	acctcgaaca	gctcccaga	aaaaacaaaa	ttgcagtcag	2640
aacatggcta	ttcgagaatg	gagaggagag	agttactaga	atggcgatca	gtggtgatga	2700
ctgtgtagtt	aaaccgctgg	acgacaggtt	cgcaacggca	ctccacttcc	tcaatgccat	2760
gtcgaaagtt	agaaaggata	tccaggaatg	gaaaccctcg	catgggtggc	acgattggca	2820
acaggtcccc	ttctgttcca	atcacttcca	ggaaattgtg	atgaaagacg	gaaggagtat	2880
tgctgtgccg	tgcaaggggc	aggacgagct	tattggaagg	gcgcgcatct	ccccaggtgc	2940
tggatggaat	gtgaaagaca	cagcatgcct	ggccaaagcg	tatgcacaaa	tgtggctgct	3000
tttgtacttc	catcggagag	acctacgcct	tatggcaaac	gccatttgca	gtgcagtacc	3060
tgtggactgg	gtgccaccg	gtagaacttc	ctggtcaata	catagcaagg	gcgaatggat	3120
gaccactgaa	gatatgctgc	aagtatggaa	cagggtatgg	atagaggaga	atgagtggat	3180
gatggataag	acaccaataa	catcctggac	agacgttccg	tacgtaggaa	agcgcgagga	3240
catctggtgt	ggcagtctaa	tcggaacgcg	aagcagagcc	acttgggcag	agaatatcca	3300
tgcggtata	aaccaagtta	gagctattat	agggaaagag	aattatgtgg	attatatgac	3360
ttcacttaaa	agatatgagg	atgtattgat	ccaagaagac	agagtcatt		3409

REIVINDICACIONES

1. Método para generar *in vitro* un virus de ARN positivo monocatenario atenuado de un receptor humano o animal que comprende las siguientes etapas:

etapa I) re-codificar el genoma vírico de un virus ARN infeccioso sustituyendo aleatoriamente una parte de los codones de nucleótidos de todo el genoma vírico de dicho virus ARN infeccioso por otro codón de nucleótidos que codifica el mismo aminoácido, con la condición de que:

- i) el número y la posición de los codones de nucleótidos raros presentes en dicho genoma vírico no se modifiquen, siendo dichos codones de nucleótidos raros CGU, CGC, CGA, CGG, UCG, CCG, GCG y ACG; y
- ii) las regiones de dicho genoma vírico que están implicadas con la estructura secundaria del ARN no se modifiquen; y,

etapa II) generar un virus de ARN atenuado mediante:

sub-etapa II.a) introducción de un promotor de la ARN polimerasa dependiente de ADN en la posición 5' y opcionalmente un terminador y una secuencia de poliadenilación de ARN en la posición 3' del genoma vírico recodificado como se obtiene en la etapa I);

sub-etapa II.b) amplificación del genoma vírico recodificado como se prepara en la sub-etapa a) incluyendo dicho promotor y opcionalmente dicho terminador y secuencia de poliadenilación de ARN, en al menos 2, preferentemente al menos 3, 4, 5 o 6 fragmentos de ADNc solapantes;

sub-etapa II.c) transfección de dichos fragmentos de ADNc en una célula hospedadora;

sub-etapa II.d) incubación de dicha célula hospedadora de la sub-etapa c); y

sub-etapa II.e) recuperación del virus de ARN atenuado de dicha célula hospedadora incubada.

2. Método de acuerdo con la reivindicación 1, en donde en la etapa I), del 1 al 20 % de los codones de nucleótidos de todo el genoma vírico de dicho virus de ARN infeccioso se sustituyen por otro codón de nucleótidos que codifica el mismo aminoácido.

3. Método de acuerdo con la reivindicación 1 o 2, en donde la etapa I) se realiza mediante:

- determinando la secuencia de aminoácidos codificada por todo el genoma vírico del virus de ARN infeccioso y determinar cada codón de nucleótido que codifica cada aminoácido; y
- sustituyendo del 1 al 20 % del codón de nucleótidos del genoma vírico que codifica un aminoácido de la tabla 1, por un codón de nucleótido diferente que codifica el mismo aminoácido como se especifica en la siguiente tabla:

Aminoácido	Codón de nucleótido
Ala, A	GCU, GCC, GCA
Arg/R	AGA, AGG
Asn/N	AAU, AAC
Asp/D	GAU, GAC
Cys/C	UGU, UGC
Gln/Q	CAA, CAG
Glu/E	GAA, GAG
Gly/G	GGU, GGC, GGA
His/H	CAU, CAC
Ile/I	AUU, AUC, AUA
Leu/L	UUA, UUG, CUU, CUC, CUA, CUG
Lys/K	AAA, AAG
Phe/F	UUU, UUC

(continuación)

Aminoácido	Codón de nucleótido
Pro/P	CCU, CCC, CCA
Ser/S	UCU, UCC, UCA, AGU, AGC
Thr/T	ACU, ACC, ACA
Tyr/Y	UAU, UAC
Val/V	GUU, GUC, GUA, GUG

4. El método de acuerdo con una cualquiera de las reivindicaciones **1 a 3**, en donde dicho virus de ARN positivo monocatenario se selecciona del grupo que consiste en flavivirus, alfavirus y enterovirus.
- 5
5. El método de acuerdo con una cualquiera de las reivindicaciones **1 a 4**, en donde, en la etapa II.a),
- dicho promotor de la ARN polimerasa dependiente de ADN en la posición 5' es el promotor del citomegalovirus humano (pCMV); y/o
- 10
- dicha secuencia terminador y de poliadenilación de ARN opcional es, respectivamente, la ribozima de la hepatitis delta y la señal de poliadenilación del virus de simio 40 (HDR/SV40pA).
6. El método de acuerdo con una de las reivindicaciones **1 a 5**, en donde:
- 15
- la sub-etapa II.c) es una etapa de transfección directa de los fragmentos de ADNc obtenidos en la sub-etapa b) como tal, y
 - dicha sub-etapa II.c) se produce directamente después de la sub-etapa II.b).
7. El método de acuerdo con una cualquiera de las reivindicaciones **1 a 5**, en donde la sub-etapa II.c) es una etapa de transfección de plásmidos o vectores que comprenden un fragmento de ADNc obtenido en la sub-etapa II.b), en donde cada fragmento de ADNc está en un plásmido o vector individual y separado.
- 20
8. El método de acuerdo con una cualquiera de las reivindicaciones **1 a 7**, en donde los fragmentos de ADNc transfectados de la sub-etapa II.c) se recombinan espontáneamente en las células hospedadoras durante la sub-etapa II.d) de incubación.
- 25
9. El método de acuerdo con una cualquiera de las reivindicaciones **1 a 8**, en donde dicho virus es el virus Chikungunya y se realiza dicha etapa I) de recodificación:
- 30
- en la región que codifica la proteína no estructural nsP1, por el casete recodificado representado en SEQ ID NO: 63;
 - en la región que codifica la proteína no estructural nsP4, por el recodificado representado en SEQ ID NO: 64; y
 - en la región que codifica la región que se solapa con las proteínas estructurales E2 y E1, por el casete recodificado representado en SEQ ID NO: 65.
- 35
10. El método de acuerdo con una cualquiera de las reivindicaciones **1 a 8**, en donde dicho virus es el virus de la encefalitis transmitida por garrapatas y dicha etapa I) de la etapa de recodificación se realiza en la región genómica NS5, por el casete recodificado representado en SEQ ID NO: 66.
- 40
11. El método de acuerdo con una cualquiera de las reivindicaciones **1 a 8**, en donde dicho virus es el virus de la encefalitis japonesa y dicha etapa I) de la etapa de recodificación se realiza en el marco de lectura abierto (ORF) completo, desde el comienzo de PrM hasta el final de la región genómica NS5 por al menos un casete recodificado seleccionado del grupo que consiste en SEQ ID NO: 67; SEQ ID NO: 68; SEQ ID NO: 69; SEQ ID NO: 70; SEQ ID NO: 71; y SEQ ID NO:72.

FIGURAS

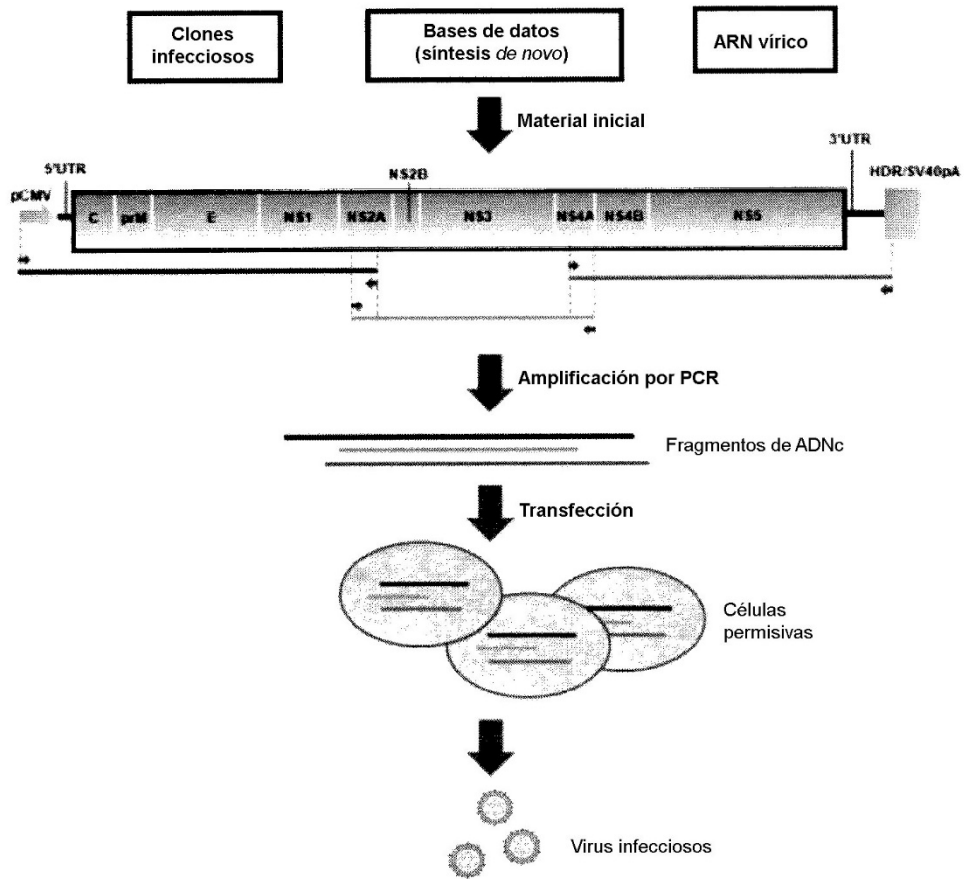


Figura 1

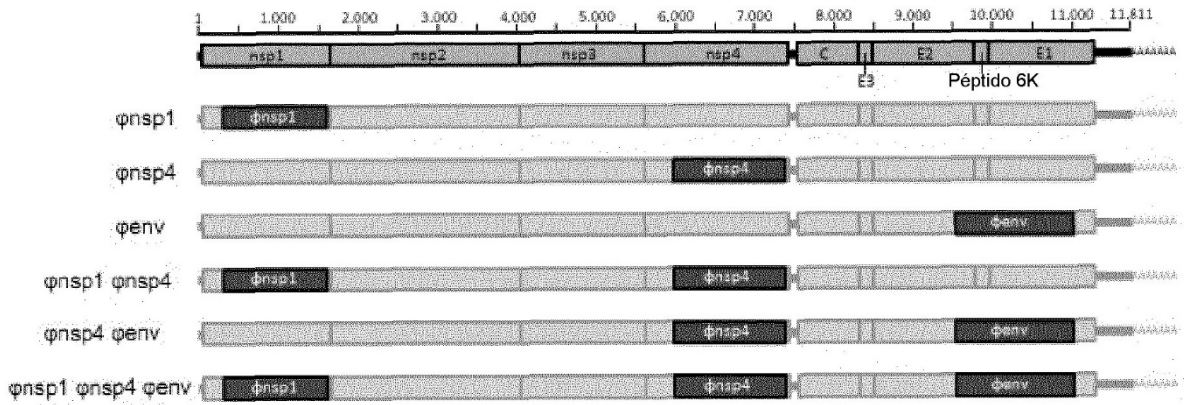


Figura 2

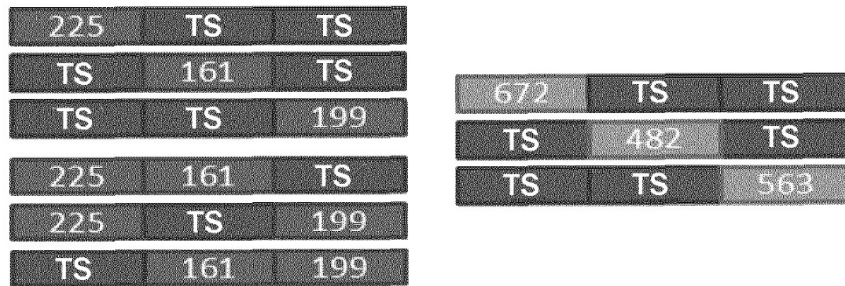


Figura 3

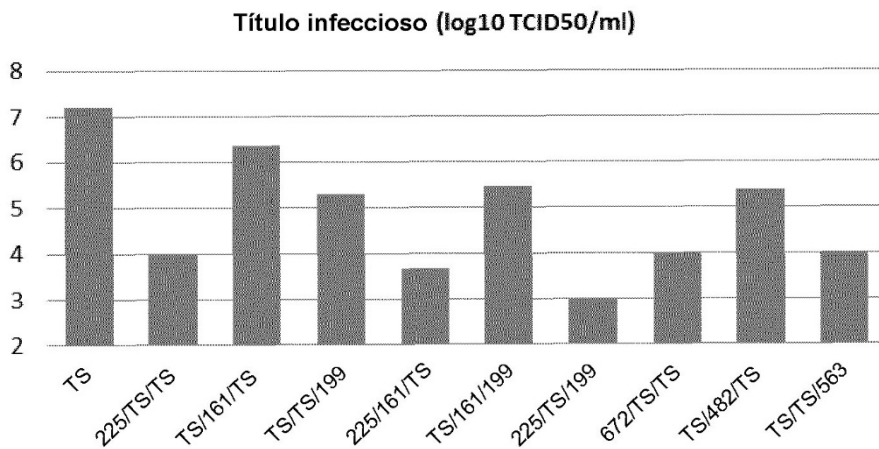


Figura 4