



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



① Número de publicación: **2 170 622**

② Número de solicitud: 009902673

⑤ Int. Cl.7: **C12N 7/04**  
C12N 15/50  
C12N 15/86  
A61K 39/12  
A61K 39/295

⑫

PATENTE DE INVENCION

B1

⑫ Fecha de presentación: **03.12.1999**

⑬ Fecha de publicación de la solicitud: **01.08.2002**

Fecha de la concesión: **14.04.2004**

⑭ Fecha de anuncio de la concesión: **16.05.2004**

⑮ Fecha de publicación del folleto de la patente:  
**16.05.2004**

⑰ Titular/es:  
**Consejo Superior de Investigaciones Científicas  
c/ Serrano 117  
28006 Madrid, ES**

⑱ Inventor/es:  
**Enjuanes Sánchez, Luis; Almazán Toral, F.;  
González Martínez, J.Manuel; Izeta Permisan, A.;  
Penzes, Zoltan; Alonso Villanueva, Sara;  
Sánchez Sánchez, Carlos; Solá Gurpegui, I.;  
Sánchez Morgado, J. Manuel; Calvo Alcocer, E.;  
Ortego Alonso, F.J.; Escors Murugarren, D.  
Barrado Guerrero, Patricio; Riquelme Gabriel, C. y  
Plana Durán, Juan**

⑲ Agente: **No consta**

⑳ Título: **Clones y vectores infectivos derivados de coronavirus y sus aplicaciones.**

㉑ Resumen:

Clones y vectores infectivos derivados de coronavirus y sus aplicaciones.

El clon infectivo derivado de un coronavirus comprende un cDNA que codifica el gRNA de un coronavirus clonado bajo una secuencia promotora de la transcripción. El vector viral recombinante comprende un clon infectivo modificado para contener un ácido nucleico heterólogo insertado en dicho clon infectivo bajo condiciones que permiten la expresión de dicho ácido nucleico heterólogo. Los clones y vectores infectivos son útiles tanto en investigación básica como aplicada, en el desarrollo de sistemas de expresión eficientes de productos de interés (proteínas, enzimas, anticuerpos, etc.), vectores vacunales y terapia génica.

ES 2 170 622 B1

Aviso: Se puede realizar consulta prevista por el art. 37.3.8 LP.

## DESCRIPCION

Clones y vectores infectivos derivados de coronavirus y sus aplicaciones.

**Campo de la invención**

Esta invención se refiere a unos clones infectivos que comprenden el DNA complementario (cDNA) al RNA genómico de un coronavirus, a vectores que comprenden dichos clones infectivos, a procedimientos para la obtención de dichos clones infectivos y vectores y a sus aplicaciones.

**Antecedentes de la invención**

Los avances en la tecnología del DNA recombinante han conducido al progreso en el desarrollo de transferencia de genes entre organismos. Actualmente se están dedicando numerosos esfuerzos para intentar producir productos químicos, farmacéuticos y biológicos de interés económico y comercial mediante el empleo de técnicas de transferencia de genes.

Uno de los elementos clave en la manipulación genética de células, tanto procarióticas como eucarióticas, es el desarrollo de vectores y sistemas vector-hospedador. En general, un vector es una molécula de ácido nucleico capaz de introducir un ácido nucleico heterólogo en una célula hospedadora, mientras que un sistema vector-hospedador puede definirse como una célula hospedadora que porta un vector y permite la replicación y expresión de la información genética incluida en este.

Se han desarrollado vectores a partir de virus tanto con genoma DNA como RNA. Los vectores virales derivados de virus DNA que se replican en el núcleo de la célula hospedadora presentan el inconveniente de poder llegar a integrarse en el genoma de dicha célula, por lo que, en general, no son muy seguros. Por el contrario, los vectores virales derivados de virus RNA que se replican en el citoplasma de la célula hospedadora son más seguros que los basados en virus DNA ya que la replicación transcurre vía RNA fuera del núcleo, con lo que las posibilidades de que se integren en el genoma de la célula hospedadora son muy pequeñas.

Se han obtenido clones de cDNA a partir de virus RNA de cadena sencilla con polaridad positiva [ss-RNA(+)], por ejemplo, picornavirus [Racaniello & Baltimore, 1981]; bromovirus [Ahlquist et al., 1984]; alfavirus, género que incluye el virus Sindbis, el virus del bosque de Semliki (SFV) y el virus de la encefalitis equina venezolana (VEE) [Rice et al., 1987; Liljeström and Garoff, 1991; Frolov et al., 1996; Smerdou and Liljeström, 1999]; flavivirus y pestivirus [Rice and Strauss, 1981; Lai et al., 1991; Rice et al., 1989]; y de virus de la familia *Astroviridae* [Geigenmuller et al., 1997]. Asimismo, se han desarrollado vectores para la expresión de genes heterólogos a partir de clones de DNA complementario al genoma de virus ssRNA(+), por ejemplo, alfavirus, entre los que se incluyen el virus Sindbis, el virus del bosque de Semliki (SFV) y el virus de la encefalitis equina venezolana (VEE) [Frolov et al., 1996; Liljeström, 1994; Pushko et al., 1997].

Aunque se han desarrollado diversos vectores virales para replicar y expresar ácidos nucleicos heterólogos en células hospedadoras, la mayoría de los vectores conocidos de expresión de genes heterólogos presentan inconvenientes relacionados con su especificidad de especie y órgano diana y con su limitada capacidad de clonaje, lo que restringe sus posibilidades de uso tanto en investigación básica como en investigación aplicada.

Por tanto, sigue existiendo la necesidad de desarrollar nuevos vectores de expresión de genes heterólogos que superen los inconvenientes mencionados. En particular, sería muy adecuado disponer de unos vectores de expresión de genes heterólogos que tuvieran una elevada seguridad y capacidad de clonaje y pudieran ser diseñados de forma que se pudiera controlar fácilmente su especificidad de especie y su tropismo.

La presente invención proporciona una solución al problema existente que comprende el desarrollo de unos clones infectivos de cDNA, derivados de coronavirus, así como unos vectores construidos a partir de dichos clones infectivos que además incluyen secuencias de ácidos nucleicos heterólogos insertadas en dichos clones. Los clones y vectores infectivos proporcionados por esta invención tienen numerosas aplicaciones tanto en investigación básica y aplicada así como una elevada capacidad de clonaje y pueden ser diseñados de forma que se pueda controlar fácilmente su especificidad de especie y su tropismo.

Los coronavirus son los virus ssRNA(+) que presentan el genoma más grande conocido para un virus RNA, con una longitud comprendida entre 27,6 y 31 kilobases (kb) [Siddell, 1995; Lai & Cavanagh, 1997;

Enjuanes et al., 1998]. Durante la infección por coronavirus se produce la replicación del RNA genómico (gRNA) y la síntesis de un conjunto de RNAs subgenómicos (sgRNA) de polaridad positiva y negativa [Sethna et al., 1989; Sawicki and Sawicki, 1990; van der Most & Spaan, 1995]. La síntesis de los sgRNAs es un proceso dependiente de RNA que ocurre en el citoplasma de la célula infectada aunque su mecanismo preciso todavía no se conoce exactamente.

Aunque se ha descrito la construcción de cDNAs que codifican genomas defectivos interferentes (DI) [mutantes de delección que requieren la presencia de un virus complementador para su replicación y transcripción] de algunos coronavirus, tales como el virus de la hepatitis murina (MHV), el virus de la bronquitis infecciosa (IBV), el coronavirus bovino (BCV) [Chang et al., 1994] y el virus de la gastroenteritis porcina (VGPT) [Solicitud de Patente Española P9600620] [Méndez et al., 1996; Izeta et al., 1999; Sánchez et al., 1999] no se ha descrito la construcción de un clon de cDNA que codifique un genoma completo de un coronavirus. El gran tamaño del genoma de los coronavirus ha sido un obstáculo importante para obtener dicho clon de cDNA.

Se describe ahora el desarrollo de un método que permite obtener por primera vez en la historia de los coronavirus, un clon cDNA infectivo que codifica el genoma de un coronavirus.

### Compendio de la invención

Se ha desarrollado un nuevo vector o sistema de expresión de ácidos nucleicos heterólogos basado en un coronavirus generado a partir de un clon cDNA infectivo que codifica el RNA genómico (gRNA) de un coronavirus. En una realización particular de esta invención, el coronavirus es el virus de la gastroenteritis porcina transmisible (VGPT).

El nuevo sistema de expresión puede ser utilizado en investigación básica o aplicada, por ejemplo, para obtener productos de interés (proteínas, enzimas, anticuerpos, etc.), como vector vacunal o en terapia génica tanto en humanos como en animales. El coronavirus infectivo obtenido a partir del clon infectivo de cDNA puede ser manipulado por técnicas convencionales de ingeniería genética con lo que se podrán introducir nuevos genes en el genoma del coronavirus y expresar estos genes de una manera específica de tejido y especie para inducir una respuesta inmune o para terapia génica. Además, se ha optimizado la expresión mediante la selección de nuevas secuencias reguladoras de la transcripción (TRS) que permiten incrementar los niveles de expresión más de 100 veces.

Los vectores derivados de coronavirus, en particular, del VGPT presentan varias ventajas para la inducción de inmunidad en mucosas con respecto a otros sistemas de expresión que no replican en estas: (i) el VGPT infecta las mucosas entéricas y respiratorias [Enjuanes and Van der Zeijst, 1995] es decir, los sitios mas adecuados para la inducción de inmunidad secretoria; (ii) su tropismo puede ser controlado mediante la modificación del gen S (spike) [Ballesteros et al., 1997]; (iii) se dispone de aislados no patógenos para el desarrollo de sistemas de expresión dependientes de virus complementador [Sánchez et al., 1992]; y, (iv) los coronavirus son virus RNA citoplásmicos que replican sin pasar por una etapa intermediaria de DNA [Lai and Cavanagh, 1997] haciendo prácticamente imposible su integración en el cromosoma celular.

El procedimiento que ha hecho posible recuperar un coronavirus infectivo a partir de un cDNA que codifica el gRNA de un coronavirus, incluye las siguientes estrategias:

- (i) la expresión del RNA del coronavirus bajo el control de un promotor apropiado;
- (ii) el clonaje del genoma del coronavirus en cromosomas bacterianos artificiales (BACs);
- (iii) la identificación de las secuencias de cDNA del coronavirus tóxicas directa o indirectamente para la bacteria;
- (iv) la identificación del orden preciso de ensamblamiento de los elementos que componen el cDNA que codifica un RNA infectivo de coronavirus (promotores, secuencias de terminación de transcripción, secuencias de poliadenilación, ribozimas, etc.); y
- (v) la identificación de un grupo de tecnologías y procesos (condiciones para el crecimiento de los BACs, modificaciones al proceso de purificación de DNA de BAC, técnicas de transformación, etc.) que combinados entre sí permiten el rescate eficiente de un coronavirus infectivo a partir de un cDNA.

El promotor juega un papel importante para aumentar la expresión del RNA viral en el núcleo, donde es sintetizado, para ser transportado posteriormente al citoplasma.

El empleo de BACs constituye uno de los puntos clave del procedimiento de la invención. Como es conocido, a menudo, el clonaje de secuencias eucarióticas en plásmidos bacterianos no es posible debido a la toxicidad para la bacteria de las secuencias exógenas. En estos casos, la bacteria suele eliminar la toxicidad modificando las secuencias introducidas. Sin embargo, en la estrategia seguida en este caso, para evitar la posible toxicidad de las secuencias virales, se realizaron los clonajes necesarios para obtener un cDNA completo del coronavirus en BACS. Estos plásmidos aparecen solo en una copia o máximo dos por célula, limitando considerablemente su toxicidad y reduciendo las posibilidades de recombinación interplásmidos.

Mediante la identificación de las secuencias de cDNA del coronavirus tóxicas para la bacteria se puede completar la construcción del cDNA que codifica el genoma completo de un coronavirus con la excepción de las secuencias tóxicas, las cuales se añaden en el último paso de la construcción del genoma completo, es decir, justo antes de la transfección en células eucarióticas, con lo que se evita su modificación por parte de la bacteria.

Un objeto de esta invención lo constituye un clon infectivo derivado de un coronavirus que comprende un cDNA que codifica el gRNA de un coronavirus, y su procedimiento de obtención.

Un objeto adicional de esta invención lo constituye un vector viral recombinante que comprende dicho clon infectivo y una secuencia de ácido nucleico heterólogo insertado dentro de dicho clon infectivo.

Un objeto adicional de esta invención lo constituye un método para producir un coronavirus recombinante que comprende la introducción de dicho clon infectivo en una célula hospedadora, el cultivo de la célula transformada en condiciones que permiten la replicación del clon infectivo y la producción del coronavirus recombinante, y recuperar el coronavirus recombinante del cultivo.

Otro objeto adicional de esta invención lo constituye un método para producir un coronavirus recombinante modificado que comprende introducir el vector viral recombinante en una célula hospedadora, cultivar ésta en condiciones que permiten la replicación del vector viral y la producción del coronavirus recombinante modificado, y recuperar el coronavirus recombinante modificado del cultivo.

Otro objeto adicional de esta invención lo constituye un método para producir un producto de interés que comprende cultivar una célula hospedadora que contiene dicho vector viral recombinante en condiciones que permiten la expresión de la secuencia de DNA heteróloga.

Las células que contienen los clones infectivos o el vector viral recombinante mencionados constituyen otro objeto adicional de la presente invención.

Otro objeto adicional de esta invención lo constituyen unas vacunas inductoras de la protección de animales frente a infecciones causadas por agentes infectivos. Estas vacunas comprenden los vectores infectivos que expresan, al menos, un antígeno adecuado para inducir una respuesta inmune frente a cada agente infectivo, o bien, al menos, un anticuerpo que proporciona protección contra dicho agente infectivo, junto con un excipiente farmacéuticamente aceptable. Las vacunas pueden ser mono-, o multivalentes dependiendo de si los vectores expresan uno o más antígenos capaces de inducir una respuesta inmune frente a uno o más agentes infecciosos, o, alternativamente, uno o más anticuerpos que proporcionan protección contra uno o más agentes infecciosos.

Otro objeto proporcionado por esta invención comprende un método de inmunización de animales que consiste en la administración de dicha vacuna.

#### Breve descripción de las figuras

La Figura 1 muestra la construcción de un clon cDNA que codifica un RNA infectivo de VGPT. La Figura 1A muestra la estructura genética del VGPT indicando mediante letras y números los nombres de los genes [1a, 1b, S, 3a, 3b, E, M, N y 7]. La Figura 1B muestra la estrategia de clonaje del cDNA que consistió en completar el genoma DI-C. Se indican las deleciones  $\Delta A1$ ,  $\Delta 2$ , y  $\Delta 3$  que se han completado para restablecer la longitud completa del cDNA. Los números situados debajo de la estructura del genoma DI-C indican los nucleótidos que flanquean cada deleción en dicho genoma DI-C. La Figura 1C muestra los cuatro fragmentos de cDNA construidos para completar la deleción  $\Delta 1$  y la posición de los principales sitios de restricción utilizados durante el ensamblaje. La inserción del fragmento  $\Delta 1$  produjo un aumento en la toxicidad del cDNA.

La Figura 2 muestra la estructura del plásmido pBeloBAC [Wang et al., 1997] utilizado en el clonaje del

## ES 2 170 622 B1

cDNA infectivo de VGPT. El plásmido pBeloBAC fue proporcionado por H. Shizuya y M. Simon (California Institute of Technology) e incluye 7.507 pares de bases (pb) que contienen el origen de replicación del factor F de *E. coli* (*oriS*), los genes necesarios para mantener una copia única del plásmido por célula (*parA*, *parB*, *parC* y *repE*) y el gen de resistencia a cloranfenicol (CM<sup>r</sup>). Se indican las posiciones de los promotores de T7 y SP6 y de los sitios de restricción únicos. CosN: sitio cosn de lambda para facilitar la construcción del plásmido pBAC; lac Z: gen de la  $\beta$ -galactosidasa. También se indica la secuencia loxP utilizada durante la generación del plásmido.

La Figura 3 muestra la estructura de los plásmidos básicos utilizados en la construcción del cDNA de VGPT. El plásmido pBAC-TcDNA<sup>- $\Delta$ ClaI</sup> contiene toda la información del RNA de VGPT excepto un fragmento ClaI-ClaI de 5.198 pb. El cDNA se clonó bajo el promotor de expresión inmediatamente temprano (IE) de citomegalovirus (CMV) y está flanqueado en el extremo 3' por una cola poli(A) con 24 residuos de A, la ribozima del virus de la hepatitis delta (HDV), y las secuencias de terminación y poliadenilación de la hormona de crecimiento bovina (BGH). El plásmido pBAC-B+C+D5' contiene el fragmento ClaI-ClaI requerido para completar el pBAC-TcDNA<sup>- $\Delta$ ClaI</sup> hasta un cDNA de longitud completa. El plásmido pBAC-TcDNA<sup>FL</sup> contiene el cDNA de longitud completa de VGPT. SAP: fosfatasa alcalina.

La Figura 4 muestra las diferencias en la secuencia de nucleótidos del gen S de los clones de VGPT PUR46-MAD (MAD) y C11. Los números indican las posiciones de los nucleótidos sustituidos, considerando como nucleótido uno de cada gen la A del codón de iniciación. Las letras situadas dentro de las barras indican el nucleótido correspondiente en la posición indicada. Las letras situadas debajo de las barras indican las sustituciones de aminoácido (aa) codificadas por los nucleótidos que están alrededor de la posición indicada.  $\Delta$ 6 nt indica una delección de 6 nucleótidos. La flecha indica la posición del codón de terminación del gen S.

La Figura 5 muestra la estrategia seguida para rescatar el VGPT infectivo a partir del cDNA de longitud completa de VGPT. El plásmido pBAC-TcDNA<sup>FL</sup> se transfectó a células ST (células de testículo de cerdo) y a las 48 h post-transfección el sobrenadante se utilizó para infectar nuevas células ST. El pase de virus se realizó a los tiempos indicados. En cada pase, se recogieron alícuotas de sobrenadante y de monocapa celular para la titulación de virus y el aislamiento de RNA para análisis por RT-PCR, respectivamente. vgRNA: RNA viral de longitud completa.

La Figura 6 muestra el efecto citopático (CPE) producido por el cDNA de VGPT en las células ST transfectadas. Se muestra la ausencia de CPE en células ST no transfectadas (control) (Figura 6A) y el CPE observado a las 14 y 20 h post-transfección con pBAC-TcDNA<sup>FL</sup> en células ST (Figuras 6B y 6C, respectivamente).

La Figura 7 muestra la evolución del título viral con el pase. Se muestra una gráfica en la que se representa el título viral en el sobrenadante de dos series de monocapas celulares (1 y 2) a diferentes pases después de la transfección con pBAC-TcDNA<sup>FL</sup>. block 1 y 2, se refieren a células ST no transfectadas. TcDNA 1 y 2, se refieren a células ST transfectadas con pBAC-TcDNA<sup>FL</sup>.

La Figura 8 muestra los resultados del análisis de la secuencia del virus recuperado después de transfectar células ST con pBAC-TcDNA<sup>FL</sup>. La estructura del genoma de VGPT se indica en la parte superior de la figura. Asimismo, se indican las diferencias en la secuencia de nucleótidos (marcadores genéticos) entre el virus recuperado a partir del plásmido pBAC-TcDNA<sup>FL</sup> (TcDNA), y los clones C8 y C11 de VGPT. La posición de las diferencias entre los nucleótidos se indica mediante los números situados sobre la barra. Las secuencias del cDNA del virus TcDNA y del clon C11 se determinaron por secuenciación de los fragmentos obtenidos por RT-PCR [transcripción inversa-reacción en cadena de la polimerasa]. La secuencia de cDNA del clon C8 del aislado PUR46-MAD de VGPT, identificada como SEC. ID. N°.: 1, se recoge en la LISTA DE SECUENCIAS. Se muestran los patrones de restricción con ClaI y DraIII de los fragmentos obtenidos por RT-PCR que incluyen los nucleótidos 18.997 y 20.990 de los virus TcDNA y C8. Los patrones de restricción muestran la presencia o ausencia de los sitios ClaI y DraIII en el cDNA de estos virus. El resultado de este análisis indicó que el virus TcDNA recuperado tenía la secuencia del gen S esperada para el aislado C11. MWM: marcadores de peso molecular.

La Figura 9 muestra los resultados del análisis por RT-PCR del virus recuperado. El RNA viral se expresó bajo el control del promotor de CMV reconocido por la polimerasa celular *pol II*. En principio, este RNA podría sufrir "splicing" durante su transporte al citoplasma. Para estudiar si éste era el caso, se determinaron los sitios del RNA con una alta probabilidad de splicing utilizando un programa de predicción de sitios de splicing en secuencias de DNA humano (version 2.1.5.94, Department of Cell Biology, Baylor College of Medicine) [Solovyev et al., 1994]. El sitio potencial de

splicing con máxima probabilidad de corte tenía el sitio donador en el nt 7.243 y el aceptor en el nt 7.570 (Figura 9A). Para estudiar si este dominio había sufrido “*splicing*”, se preparó un fragmento de RT-PCR flanqueado por los nt 7.078 y 7.802 (Figura 9B) a partir de RNA de los pases 0 y 2 de cultivos no transfectados (control), o de células ST transfectadas con el TcDNA con el fragmento ClaI en orientación reversa (TcDNA<sup>FL(-Δ*ClaI*)RS</sup>), o en la orientación correcta (TcDNA<sup>FL</sup>), y se analizaron los productos resultantes de la RT-PCR en geles de agarosa. Los resultados obtenidos se muestran en las Figuras 9C (pase 0) y 9D (pase 2).

La Figura 10 muestra los resultados del análisis por inmunofluorescencia del virus producido en cultivos de células ST transfectadas con el TcDNA. Se realizó tinción para inmunofluorescencia con anticuerpos específicos para el aislado VGPT PUR46-MAD, y para el virus recuperado después de la transfección con el plásmido pBAC-TcDNA<sup>FL</sup>. Para ello, se emplearon anticuerpos monoclonales específicos para el VGPT que se unen a ambos aislados o sólo al PUR46-MAD [Sánchez, et al., 1990]. El resultado confirmó que el virus TcDNA tenía la antigenicidad esperada. El antisuero policlonal específico de VGPT se unió a ambos virus, pero no a los cultivos no infectados y sólo se unieron al virus TcDNA los anticuerpos monoclonales esperados, específicos para las proteínas S (ID.B12 y 6A.C3), M (3B.B3) y N (3B.D8) [Sánchez et al., 1999].

La Figura 11 muestra la expresión de GUS bajo diferentes secuencias reguladoras de la transcripción (TRSs) que varían la región flanqueante 5' de la secuencia intergénica (CTAAAC, IG). El minigenoma M39 se clonó bajo el control del promotor de CMV. En el interior de este minigenoma se insertó una secuencia múltiple de clonaje [PL1, 5'-CCTAGGATTTAAATCCTAAGG-3' (SEC. ID. N°.: 2)] y la unidad de transcripción formada por las secuencias reguladoras de la transcripción seleccionadas (TRS), otra secuencia múltiple de clonaje [PL2, 5'-GCGGCCGCGCCGCGAGGCCTGTCGAC-3' (SEC. ID. N°.: 3) o PL3, 5'-GTCGAC-3' (SEC. ID. N°.: 4)], secuencias con la estructura de un dominio de Kozak (K), el gen de la β-glucuronidasa (GUS), y otro sitio múltiple de clonaje [PL4, 5'-GCTAGCCCAGCCGCGCGGTACC-3' (SEC. ID. N°.: 5)]. Estas secuencias estaban flanqueadas en el extremo 3' por las secuencias 3' del minigenoma M39 (indicadas por un cuadrado gris), la ribozima de HDV, y las secuencias de terminación y poliadenilación de la BGH. Las TRSs tenían distinto número (0, -3, -8, y -88) de nucleótidos en el extremo 5' de la secuencia IG (CTAAAC) y procedían de los genes N, S, o M, tal como se indica. Se transfectaron células ST con los distintos plásmidos, se infectaron con el virus complementador (PUR46-MAD), y los sobrenadantes se sometieron a 6 pases. La actividad GUS en las células infectadas se determinó mediante el protocolo descrito por Izeta [Izeta y col., 1999]. Los resultados obtenidos al relacionar la actividad GUS, expresada como unidades luminométricas relativas por millón de células, con el número de pase se recogen en la Figura 11B.

La Figura 12 muestra la expresión de GUS bajo diferentes TRSs que varían en la región flanqueante 3' de la secuencia IG (véase la Figura 11A). Utilizando esta unidad de transcripción con la región 5' flanqueante correspondiente a los -88 nt del gen N de VGPT más la secuencia IG (CTAAAC), se modificaron las secuencias 3' flanqueantes. Estas secuencias correspondían a las de los distintos genes de VGPT (S, 3a, 3b, E, M, N, y 7), tal como se indica en la Figura 12A. En dos casos, las secuencias 3' se reemplazaron por otras que contenían un sitio de restricción (SalI) y una secuencia de Kozak optimizada (Kz), o por una secuencia idéntica a la que sigue a la primera secuencia IG situada a continuación del líder del genoma viral. La actividad GUS en las células infectadas se determinó mediante el protocolo descrito previamente [Izeta et al., 1999]. cL12, indica una secuencia de 12 nucleótidos idéntica a la del extremo 3' de la secuencia “*leader*” del genoma del virus TGEV (ver la secuencia del virus indicada al final). Los resultados obtenidos al relacionar la expresión de GUS con el número de pase se recogen en la Figura 12B.

La Figura 13 muestra el efecto del sitio de inserción del módulo de expresión en el minigenoma sobre los niveles de expresión de GUS. La unidad de transcripción de GUS que contiene -88 nt del gen N flanqueando el extremo 5' de la secuencia IG (CTAAAC), y las secuencias Kz flanqueando el extremo 3' (véase la Figura 12A), se insertó en cuatro sitios únicos de restricción en el minigenoma M39 (Figura 13A) para determinar si todos estos sitios eran igualmente permisivos para la expresión del gen heterólogo. Se transfectaron células ST con estos plásmidos y se infectaron con el virus complementador (PUR46-MAD). La actividad GUS en las células infectadas se determinó en el pase 0 (PO) siguiendo el protocolo descrito previamente [Izeta y col., 1999]. Los resultados obtenidos se recogen en la Figura 13B.

## Descripción detallada de la invención

La invención proporciona un clon infectivo derivado de un coronavirus, en adelante, clon infectivo de la invención, que comprende un DNA complementario (cDNA) que codifica el RNA genómico (gRNA)

## ES 2 170 622 B1

de un coronavirus clonado bajo una secuencia promotora de la transcripción.

La expresión “clon infectivo” así como “clon cDNA infectivo” se utilizan indistintamente en esta descripción para referirse a un plásmido que contiene un clon de cDNA que codifica un RNA infectivo de un coronavirus. El cDNA puede ser de longitud completa, o no, si bien en este último caso mantiene toda la funcionalidad ya que codifica un RNA infectivo.

En una realización particular de esta invención, el coronavirus es un aislado de VGPT, en particular, el aislado PUR46-MAD [Sánchez et al., 1990], modificado por el reemplazamiento del gen S de este virus por el gen S del aislado VGPT clon C11 [Sánchez et al., 1999].

La secuencia promotora de la transcripción, o promotor, es una secuencia RNA situada en el extremo 5'-terminal de cada RNA mensajero (mRNA) de coronavirus, a la que se une la RNA polimerasa viral para iniciar la transcripción del RNA mensajero (mRNA). En una realización particular y preferida, el promotor es el promotor IE de CMV debido al elevado nivel de expresión obtenido usando este promotor [Dubensky et al., 1996] y a resultados previos obtenidos en nuestro laboratorio que indicaban que genomas defectivos de gran tamaño (9,7 kb y 15 kb) derivados del coronavirus VGPT expresaban RNAs que no sufrían splicing durante su transporte desde el núcleo, donde son sintetizados, hasta el citoplasma.

El clon infectivo de la invención también contiene una secuencia de terminación de transcripción y una señal de poliadenilación como la procedente del gen de la BGH. Estas secuencias de terminación tienen que estar colocadas en el extremo 3' de la cola poli (A). En una realización particular, el clon infectivo de la invención contiene una cola poli(A) de 24 residuos de A y las secuencias de terminación y poliadenilación de la BGH separadas de la cola poli(A) por la secuencia de la ribozima del HDV.

En el clon infectivo de la invención, el cDNA está integrado dentro de un plásmido. El plásmido en el que se ha insertado el cDNA infectivo del coronavirus es una molécula de DNA que posee un origen de replicación y es, por tanto, potencialmente capaz de replicarse en una célula adecuada. El plásmido utilizado es un replicón adecuado para mantener y amplificar el clon infectivo de la invención en una célula hospedadora adecuada, tal como, una bacteria, por ejemplo, *Escherichia coli*. El replicón, en general, porta un gen de resistencia a antibióticos que permite la selección de las células que lo llevan (por ejemplo, el *cat*).

En el Ejemplo 1 se describe la construcción de un clon infectivo de VGPT bajo el control del promotor IE de CMV. El extremo 3' del cDNA aparece flanqueado por una secuencia poli (A) de 24 nt, la ribozima de HDV y la secuencia de terminación de la transcripción de la BGH.

El procedimiento de obtención del clon infectivo de la invención comprende construir el cDNA que codifica el gRNA de un coronavirus y ensamblar los elementos reguladores de la transcripción.

En una realización particular, el cDNA que codifica el gRNA infectivo de un coronavirus se obtuvo a partir de un genoma DI derivado de un coronavirus que se clona como un cDNA bajo el control de un promotor apropiado en un BAC, con el fin de aumentar la estabilidad del cDNA. A continuación, se identificaron las secuencias tóxicas para la bacteria y, con el fin de eliminar esa toxicidad, se retiraron dichas secuencias tóxicas y se insertaron al final de la construcción del genoma completo, justo antes de efectuar la transfección en células eucarióticas. La progenie viral se puede reconstituir mediante transfección del plásmido BAC que contiene el genoma de coronavirus en células eucarióticas que soportan la replicación viral.

El ensamblaje de los elementos reguladores de la transcripción se realiza mediante técnicas convencionales [Maniatis et al., 1989].

El clon infectivo de la invención se puede manipular por técnicas de ingeniería genética convencionales para insertar, al menos, una secuencia de un ácido nucleico heterólogo, que codifica una determinada actividad, bajo el control del promotor que esté presente en el clon infectivo y de las secuencias reguladoras contenidas en el vector de expresión resultante.

El clon infectivo de la invención presenta numerosas aplicaciones, por ejemplo, puede ser utilizado tanto en investigación básica, por ejemplo, para estudiar el mecanismo de replicación y transcripción de los coronavirus, como en investigación aplicada, por ejemplo, en el desarrollo de sistemas de expresión eficientes de productos de interés (proteínas, enzimas, anticuerpos, etc.).

## ES 2 170 622 B1

A partir del clon infectivo de la invención se pueden transformar células apropiadas, y recuperar los viriones obtenidos que contienen el genoma completo del coronavirus. Por tanto, la invención proporciona, además, un método para producir un coronavirus recombinante que comprende la introducción de un clon infectivo de la invención en una célula hospedadora, el cultivo de dicha célula bajo condiciones que permiten la expresión y replicación del clon infectivo y la recuperación los viriones obtenidos del coronavirus recombinante que contienen el genoma infectivo del coronavirus. La introducción del clon infectivo de la invención en la célula hospedadora puede realizarse de varias maneras, por ejemplo, por transfección de la célula hospedadora con un RNA transcrito *in vitro* a partir de un clon infectivo de la invención, o por infección de la célula hospedadora con el clon cDNA infectivo de la invención. Dichas células hospedadoras que contienen el clon infectivo de la invención constituyen un objeto adicional de la presente invención.

La invención también proporciona unos vectores virales recombinantes derivados de un clon infectivo de la invención, en adelante vectores virales de la invención. Los vectores virales de la invención comprenden un clon cDNA infectivo de la invención modificado para contener un ácido nucleico heterólogo insertado en dicho clon infectivo bajo condiciones que permiten la expresión de dicho ácido nucleico heterólogo.

El término “ácido nucleico” tal como se utiliza en esta descripción incluye genes o fragmentos de genes así como, en general, cualquier molécula de DNA o RNA.

En el sentido utilizado en esta descripción, el término “heterólogo” aplicado a un ácido nucleico se refiere a una secuencia de ácido nucleico que no está presente normalmente en el clon infectivo de la invención o en el vector empleado para introducir el ácido nucleico heterólogo en una célula hospedadora.

El ácido nucleico heterólogo que puede contener el vector viral de la invención puede ser un gen o fragmento que codifica una proteína, un péptido, un epítipo o cualquier producto génico de interés (tales como anticuerpos, enzimas, etc). El ácido nucleico heterólogo se puede insertar en el clon infectivo de la invención mediante técnicas de ingeniería genética convencionales en cualquier región apropiada del cDNA, por ejemplo, después de la ORF 1b, siguiendo el codon iniciador (AUG) y en fase de lectura con ese gen; o, alternativamente, en las zonas correspondientes a otras ORFs. En la construcción del vector viral de la invención es esencial que la inserción del ácido nucleico heterólogo no interfiera ninguna de las funciones virales básicas.

El vector viral de la invención puede expresar una o más actividades. En este último caso, el vector viral incluirá tantas secuencias de ácido nucleico heterólogo como actividades se van a expresar precedidas de uno o varios promotores, o bien de un promotor y varios sitios de reconocimiento del ribosoma (IRES), o bien de varios promotores y un sitio de reconocimiento del ribosoma.

Por consiguiente, la invención proporciona un método para producir un producto de interés que comprende cultivar una célula hospedadora que contiene un vector viral de la invención bajo condiciones que permiten la expresión del ácido nucleico heterólogo y recuperar el producto de interés. Dichas células hospedadoras que contienen el vector viral de la invención constituyen un objeto adicional de la presente invención.

El vector viral de la invención puede ser diseñado de forma que se puede controlar fácilmente su especificidad de especie y su tropismo. Debido a estas características, una aplicación muy interesante de los vectores virales de la invención es su empleo en terapia génica como vector del gen de interés o como vector vacunal para inducir respuestas inmunes frente a diferentes patógenos.

La invención proporciona, además, vacunas capaces de inducir protección en un animal frente a la infección causada por un agente infeccioso que comprende (i) al menos, un vector viral de la invención que expresa, al menos, un antígeno adecuado para inducir una respuesta inmune frente a dicho agente infeccioso, o un anticuerpo que proporciona protección contra dicho agente infeccioso, junto con, opcionalmente, (ii) un excipiente farmacéuticamente aceptable.

Por “inducir protección”, en el sentido utilizado en esta descripción, debe entenderse la respuesta inmune del organismo receptor (animal a inmunizar) inducida por el vector viral de la invención, a través de los mecanismos adecuados tales como los inducidos por sustancias potenciadoras de la respuesta celular (inter-leuquinas, interferones, etc.), factores de necrosis celulares y sustancias similares que hacen que el animal quede protegido frente a infecciones causadas por agentes infecciosos.



## ES 2 170 622 B1

Bajo el término “animal” están incluidos todos los animales de cualquier especie, preferentemente mamíferos, incluido el hombre.

5 El término “agente infeccioso” en el sentido utilizado en esta descripción incluye a cualquier agente infectivo viral, bacteriano, fúngico, parasitario, u otros que puedan infectar a un animal y ocasionarle una patología.

10 En una realización particular, la vacuna proporcionada por esta invención comprende al menos un vector viral de la invención que expresa, al menos, un antígeno capaz de inducir una respuesta inmune sistémica y/o una respuesta inmune en mucosas frente a distintos agentes infecciosos que se propagan en mucosas respiratorias o entéricas. Los vectores objeto de la invención son muy adecuados para inducir inmunidad en mucosas así como inmunidad lactogénica, de especial interés en la protección de neonatos frente a infecciones del tracto intestinal.

15 En otra realización particular, la vacuna proporcionada por esta invención comprende, al menos, un vector viral de la invención que expresa, al menos, un gen que codifica para las cadenas pesada y ligera de un anticuerpo de cualquier isotipo (por ejemplo, IgG<sub>1</sub>, IgA, etc.) que proporciona protección contra un agente infeccioso.

20 La especificidad de especie se puede controlar de forma que el vector viral exprese la proteína S de la envuelta, de un coronavirus que infecta la especie deseada (humana, canina, felina, porcina, etc.), adecuada para ser reconocido por los receptores celulares de la especie correspondiente.

25 Las vacunas proporcionadas por esta invención pueden ser monovalentes o multivalentes dependiendo de si los vectores virales de la invención expresan uno o más antígenos capaces de inducir una respuesta inmune frente a uno o más agentes infecciosos o bien uno o más anticuerpos que proporcionan protección contra uno o más agentes infecciosos.

30 En una realización particular de esta invención se proporcionan vacunas monovalentes capaces de proteger al hombre, cerdos, perros y gatos contra distintos agentes infecciosos humanos, porcinos, caninos y felinos, y el tropismo se controla expresando la glicoproteína S del coronavirus con la especificidad de especie deseada.

35 Las vacunas monovalentes contra agentes infecciosos humanos pueden contener un vector que exprese un antígeno seleccionado del grupo esencialmente constituido por antígenos de patógenos humanos, especialmente agentes infecciosos humanos seleccionados entre aquellos causantes de infecciones de mucosas entéricas y/o respiratorias o bien cuya vía de entrada sea a través de dichas mucosas, por ejemplo, coronavirus humanos, rinovirus humanos, rotavirus humanos, enterovirus humanos, virus de la gripe, y, en general, cualquier patógeno infeccioso, incluyendo virus y bacterias, que infectan a humanos.

40 Las vacunas monovalentes contra agentes infecciosos porcinos pueden contener un vector que exprese un antígeno seleccionado del grupo esencialmente constituido por antígenos de los siguientes patógenos porcinos: *Actinobacillus pleuropneumoniae*, *Actinobacillus suis*, *Haemophilus parasuis*, *Parvovirus porcino*, *Leptospira*, *Escherichia coli*, *Erysipelotrix rhusiopathiae*, *Pasterella multocida*, *Bordetella bronchiseptica*, *Clostridium pp.*, *Serpulina hydrosenteriae*, *Mycoplasma hyopneumoniae*, virus de la diarrea epidémica porcina (PEDV), coronavirus respiratorio porcino, rotavirus, o contra los patógenos causantes del síndrome respiratorio y reproductivo porcino, la enfermedad de Aujeszky (Pseudorabies), influenza porcina o gastroenteritis transmisible y el agente etiológico de la rinitis atrófica y de la ileitis proliferativa.

50 Las vacunas monovalentes contra agentes infecciosos caninos pueden contener un vector de expresión que exprese un antígeno seleccionado del grupo esencialmente constituido por antígenos de los siguientes patógenos caninos: herpesvirus caninos, adenovirus canino tipos 1 y 2, parvovirus canino tipos 1 y 2, reovirus canino, rotavirus canino, coronavirus canino, virus de la parainfluenza canina, virus de la influenza canina, virus del moquillo (Distemper virus), virus de la rabia, retrovirus y calicivirus canino.

55 Las vacunas monovalentes contra agentes infecciosos felinos pueden contener un vector de expresión que exprese un antígeno seleccionado del grupo esencialmente constituido por antígenos de los siguientes patógenos felinos: calicivirus del gato, virus de la inmunodeficiencia felina, herpesvirus felinos, virus de la panleucopenia felina, reovirus felino, rotavirus felino, coronavirus felino, virus de la peritonitis infecciosa del gato, virus de la rabia, *Chlamydia psittaci* felina, y virus de la leucemia felina.

Los vectores pueden expresar un anticuerpo que proporciona protección contra un agente infeccioso,

## ES 2 170 622 B1

por ejemplo, un agente infeccioso porcino, canino o felino como los citados previamente. En una realización particular, el vector expresa el anticuerpo monoclonal recombinante identificado como 6A.C3 que neutraliza el VGPT, expresado con isotipos IgG<sub>1</sub> o IgA en el que la parte constante de la inmunoglobulina es de origen porcino o anticuerpos neutralizantes para rotavirus humanos y porcinos.

5 Como excipiente puede utilizarse un diluyente tal como suero salino fisiológico u otras soluciones salinas similares. Asimismo, estas vacunas pueden contener también un adyuvante de los habitualmente utilizados en la formulación de vacunas, tanto acuoso, tal como hidróxido de aluminio, QuilA, suspensiones de geles de alúmina y similares, como oleoso, a base de aceites minerales, glicéridos y derivados de ácido graso, y sus mezclas.

15 Estas vacunas también pueden contener sustancias potenciadoras de la respuesta celular (PRC), es decir, sustancias potenciadoras de subpoblaciones de células T helper (Th<sub>1</sub> y Th<sub>2</sub>) tales como interleuquina-1 (IL-1), IL-2, IL-4, IL-5, IL-6, IL-12, g-IFN (gamma interferón), factor de necrosis celular y sustancias similares, que podrían, teóricamente, provocar inmunidad celular en los animales vacunados. Estas sustancias PRC podrían utilizarse en formulaciones vacunales con adyuvantes acuosos u oleosos. También pueden utilizarse otro tipo de adyuvantes que modulan e inmunoestimulan la respuesta celular tales como el MDP (muramil dipéptido), ISCOM (Immuno Stimulant Complex) o liposomas.

20 La invención proporciona vacunas multivalentes capaces de prevenir y proteger animales de las infecciones causadas por distintos agentes infecciosos. Estas vacunas multivalentes pueden elaborarse a partir de vectores virales de la invención en los que se han insertado las distintas secuencias que codifican los antígenos correspondientes en el mismo vector recombinante o bien construyendo vectores recombinantes independientes que posteriormente se mezclarían para su inoculación conjunta. Por tanto, estas vacunas multivalentes comprenden un vector viral que contiene más de una secuencia de ácidos nucleicos heterólogos que codifican para más de un antígeno o, alternativamente, distintos vectores virales que expresan, cada uno de ellos, al menos un antígeno distinto.

30 Análogamente, se pueden preparar vacunas multivalentes que comprenden vectores multivalentes utilizando secuencias que codifican anticuerpos que proporcionan protección contra agentes infecciosos en lugar de secuencias que codifican los antígenos.

35 En una realización particular de esta invención se proporcionan vacunas capaces de conferir inmunidad a hombres, cerdos, perros y gatos contra distintos agentes infecciosos humanos, porcinos, caninos y felinos, respectivamente. Para ello, los vectores virales contenidos en la vacuna deben expresar distintos antígenos de los patógenos humanos, porcinos, caninos o felinos previamente mencionados.

40 Las vacunas de esta invención pueden presentarse en forma líquida o liofilizada y pueden prepararse suspendiendo los sistemas recombinantes en el excipiente. Si dichos sistemas estuvieran en forma liofilizada, el propio excipiente podría ser el reconstituyente.

45 Alternativamente, las vacunas proporcionadas por esta invención se pueden utilizar en combinación con otras vacunas convencionales, ya sea formando parte de las mismas o bien como diluyente o fracción liofilizada para diluirse con otras vacunas ya sean convencionales o recombinantes.

Las vacunas proporcionadas por esta invención pueden administrarse al animal por vía oral, nasal, subcutánea, intradérmica, intraperitoneal, intramuscular o por medio de aerosol.

50 La invención también proporciona un método para la inmunización de animales, en particular, hombres, cerdos, perros y gatos, contra uno o varios agentes infecciosos de forma simultánea, que comprende la administración por vía oral, nasal, subcutánea, intradérmica, intraperitoneal, intramuscular o por medio de aerosol (o formas combinadas de éstas) de una vacuna que contiene una cantidad inmunológicamente eficaz de un sistema recombinante proporcionado por esta invención.

55 Adicionalmente, la invención también proporciona un método para proteger a los animales recién nacidos contra agentes infecciosos que infectan a dichos animales, que consiste en la administración por vía oral, nasal, subcutánea, intradérmica, intraperitoneal, intramuscular o por medio de aerosol (o formas combinadas de éstas) a las madres antes de o durante el periodo de gestación, o a su progenie, una vacuna de las proporcionadas por esta invención.

60 La invención se ilustra mediante el siguiente ejemplo que describe de forma detallada la obtención de clones infectivos y la construcción de vectores virales de la invención. Este ejemplo no debe ser conside-

## ES 2 170 622 B1

rado como limitativo del alcance de la invención sino como ilustrativo de la misma. En dicho Ejemplo, la transformación de bacterias, el crecimiento de las bacterias, la purificación del DNA, el análisis de secuencias y el ensayo para evaluar la estabilidad de los plásmidos, se realizaron según la metodología que se describe a continuación.

### 5 *Transformación de bacterias*

Todos los plásmidos fueron electroporados en la estirpe de *E. coli* DH10B (Gibco BRL), introduciendo ligeras modificaciones a protocolos previamente descritos [Shizuya et al., 1992]. Para cada transformación  
10 mezclaron 2  $\mu$ l de la ligación y 50  $\mu$ l de bacterias competentes en cubetas de 0,2 cm (BioRad) y se electroporaron a 200  $\Omega$ , 2,5 kV y 25  $\mu$ F. A continuación, se añadió 1 ml de medio SOC [Maniatis et al., 1989] a cada transformación, se incubaron las células a 37°C durante 45 minutos, y finalmente las colonias recombinantes se detectaron en placas de LB SOC [Maniatis et al., 1989] con 12,5  $\mu$ g/ml de cloranfenicol.

### 15 *Condiciones de crecimiento de las bacterias*

Las bacterias que contenían los plásmidos originales en los que se encontraba clonado el genoma incompleto del VGPT (Figura 3) se crecieron a 37°C, mostrando cinéticas de crecimiento normales. Por otra parte, el BAC que contenía el cDNA completo se creció a 30°C con el fin de minimizar al máximo la  
20 inestabilidad. Aún así, el tamaño de las colonias era reducido y fueron necesarios periodos de incubación de hasta 24 h para conseguir tamaños de colonia normales.

### *Purificación de DNA*

Se siguió el protocolo descrito por Woo (Woo et al., 1994) con ligeras modificaciones. A partir de una única colonia se inocularon 4 l de LB con cloranfenicol (12,5  $\mu$ g/ml). Después de un periodo de incubación de 18 h a 30°C, las bacterias fueron recogidas por centrifugación a 6.000 g, y el plásmido purificado usando el kit de Maxipreparaciones de plásmidos de Qiagen de acuerdo con las recomendaciones del fabricante. Mediante este procedimiento se observó que el DNA plasmídico obtenido presentaba  
30 contaminación con DNA bacteriano. Con el fin de eliminar el DNA bacteriano contaminante, el DNA plasmídico fue purificado mediante centrifugación a 55.000 rpm durante 16 h en un gradiente de CsCl. El rendimiento obtenido fue entre 15 y 30  $\mu$ g por litro dependiendo del tamaño del plásmido.

### *Análisis de secuencias*

35 La secuenciación de DNA fue realizada en un secuenciador automático (373 DNA Sequencer, Applied Biosystems) utilizando didesoxinucleótidos marcados con fluorocromos y una polimerasa resistente a temperatura (Perkin Elmer). Los reactivos fueron obtenidos a modo de kit (ABI PRISM Dye Terminator Cycle Sequencing Ready Reaction Kit) de la compañía Applied Biosystems. El termociclador utilizado  
40 para hacer las reacciones de secuenciación fue un "GeneAmpPCR System 9600" (Perkin Elmer).

El ensamblaje de las secuencias y su comparación con la secuencia consenso del VGPT fueron realizados utilizando los programas SeqMan II y Align (DNASTAR), respectivamente. No se detectaron diferencias en relación con la secuencia consenso.

### 45 *Estabilidad de los plásmidos*

A partir de los glicerolados originales, las bacterias que contenían plásmidos pBeloBAC11 recombinantes fueron crecidas en 20 ml de LB con cloranfenicol (12,5  $\mu$ g/ml) durante 16 h a 30°C y 37°C.  
50 Este material fue considerado como pase 0. Las bacterias fueron diluidas  $10^6$  veces y crecidas a 30°C y 37°C durante 16 h. Se realizaron pases seriados durante ocho días consecutivos (cada pase representa aproximadamente 20 generaciones). El DNA plasmídico fue purificado por Miniprep a pase 0 y 8 (160 generaciones) y analizado con endonucleasas de restricción. Los dos plásmidos que contenían parte del genoma del VGPT presentaron una elevada estabilidad, mientras que el plásmido que contenía el genoma  
55 completo del VGPT mostró cierta inestabilidad después de 40 generaciones (en este punto aproximadamente el 80 % del DNA presentaba el patrón de restricción correcto).

60

## Ejemplo 1

*Construcción de un vector recombinante basado en un clon de cDNA infectivo derivado del VGPT*

5 1.1 *Generación de un cDNA infectivo del VGPT*

Con el fin de obtener un cDNA que codificaba para el genoma completo del VGPT, se partió originalmente de un cDNA que codificaba para el genoma defectivo DI-C [Méndez et al., 1996]. Este cDNA, con una longitud aproximada de un tercio del genoma del VGPT, fue clonado en el plásmido de baja copia pACNR1180 [Ruggli et al., 1996] y su secuencia determinada. El cDNA que codificaba el genoma defectivo fue rescatado eficientemente (replicado y empaquetado) con la ayuda de un virus complementador [Méndez et al., 1996; Izeta et al., 1999].

El genoma DI-C presenta tres deleciones ( $\Delta 1$ ,  $\Delta 2$  y  $\Delta 3$ ) de aproximadamente 10, 1 y 8 kilobases (kb), en las ORFs 1a, 1b, y entre los genes S y 7, respectivamente (véase la Figura 1).

La estrategia seguida para completar la secuencia de un cDNA que codificara para un genoma infectivo del VGPT fue incorporar paso a paso las secuencias delecionadas en el genoma DI-C, analizando la toxicidad en bacterias de las nuevas construcciones generadas. Este aspecto es de gran importancia ya que está ampliamente documentado en la literatura científica que plásmidos recombinantes que presentaban cDNAs de virus RNA generalmente crecían mal y eran inestables [Boyer and Haenni, 1994; Rice et al., 1989; Mandl et al., 1997].

La primera deleción en completarse fue la deleción  $\Delta 2$ , de 1 kb, de la ORF 1b, dando como resultado un plásmido recombinante estable. La secuencia que faltaba de la ORF la fue introducida clonando los fragmentos de cDNA A, B, C y D (Figura 1) [Almazan et al., 1999] de tal modo que se completara toda la información requerida para el gen de la replicasa. El plásmido recombinante obtenido era inestable en la bacteria generándose nuevos plásmidos que habían incorporado adiciones y deleciones en el fragmento B [Almazan et al., 1999]. Interesantemente, la eliminación de un fragmento de restricción ClaI-ClaI de 5.198 pb que abarcaba la región del genoma comprendida entre los nucleótidos 4.417 y 9.615 [Penzes et al., 1999] permitió la obtención de un plásmido relativamente estable en la estirpe de *E. coli* DH10B. Posteriormente, la secuencia de la deleción  $\Delta 3$  fue añadida mediante el clonaje de toda la información genética para las proteínas estructurales y las proteínas no estructurales del extremo 3' del genoma del VGPT (Figura 1).

Con el fin de incrementar la estabilidad del cDNA del VGPT se decidió subclonarlo en BAC utilizando el plásmido pBeloBAC11 [Kim et al., 1992] (véase la Figura 2). El plásmido pBeloBAC11 fue cedido generosamente por H. Shizuya y M. Simon (California Institute of Technology). El plásmido, con un tamaño de 7.507 pb, incluye el origen de replicación del factor F de *E. coli* (*oriS*), y los genes necesarios para mantener una única copia del plásmido por célula (*parA*, *parB*, *parC* and *repE*). El plásmido también presenta el gen de resistencia a cloranfenicol como marcador de selección. El cDNA fue clonado bajo el control del promotor IE de CMV debido al elevado nivel de expresión obtenido usando este promotor (Dubensky et al., 1996) y a resultados previos obtenidos en nuestro laboratorio que indicaban que genomas defectivos de gran tamaño (9,7 kb- y 15 kb) derivados del VGPT expresaban RNAs, que no sufrían splicing durante su transporte desde el núcleo, donde son sintetizados, hasta el citoplasma [Izeta et al., 1999; Penzes et al., 1999; Almazan et al., 1999]. El cDNA del VGPT generado (pBAC-TcDNA- $\Delta$ ClaI) contenía la información para los genes de la replicasa, con la excepción del fragmento ClaI de 5.198 pb delecionado, y toda la información de los genes estructurales y no estructurales. El extremo 3' del cDNA aparece flanqueado por una secuencia poli A de 24 nt, la ribozima de HDV y la secuencia de terminación de la transcripción de la BGH [Izeta et al., 1999]. Por otra parte, el fragmento ClaI necesario para generar un genoma completo del VGPT fue clonado en BAC generando el plásmido pBAC-B+C+D5', el cual contenía la región del genoma del VGPT entre los 4.310 y 9.758 [véase la Figura 3]. Ambos plásmidos fueron crecidos en la estirpe de *E. coli* DH10B y secuenciados en su totalidad. La secuencia obtenida resultó ser idéntica a la secuencia consenso del aislado PUR46-MAD del VGPT proporcionada al final de este documento, con la excepción de dos sustituciones en las posiciones de los nucleótidos 6.752 (A  $\Rightarrow$  G, silenciosa) y 18.997 (T  $\Rightarrow$  C, silenciosa) y los cambios en el gen S del PUR46 -MAD que se ha sustituido por el gen D del aislado C11 (estos cambios se indican en la Figura 4).

Además, con el fin de generar un cDNA que codificara un VGPT virulento, el gen S del aislado PUR46-MAD, que replica a elevados niveles en el tracto respiratorio ( $> 10^6$  ufp/g de tejido) y con niveles bajos en el tracto entérico ( $< 10^3$  ufp/ml), fue reemplazado completamente por el gen S del clon 11 de VGPT, en adelante C11, el cual replica con títulos elevados tanto en el tracto respiratorio ( $< 10^6$  ufp/ml) como en el entérico ( $< 10^6$  ufp/ml) [Sánchez et al., 1999]. El gen S de C11 presenta 14 nucleótidos dife-

rentes con respecto al gen S del aislado PUR46 -MAD, más una inserción de 6 nt en el extremo 5' del gen S (véase la Figura 4) [Sánchez et al., 1999]. Resultados previos en nuestro laboratorio [Sánchez et al., 1999] pusieron de manifiesto que mutantes generados por recombinación dirigida, en los cuales el gen S del aislado PUR46-MAD del VGPT fue reemplazado por el gen S del aislado entérico C11, adquirirían un tropismo entérico e incrementaban la virulencia, a diferencia del aislado natural PUR46-MAD del VGPT que replica muy poco o nada en el tracto entérico de cerdos infectados.

Se construyó un cDNA del aislado PUR46-MAD de VGPT con el gen S del aislado entérico C11 mediante clonaje del fragmento ClaI-ClaI de 5.198 pb, obtenido del plásmido pBAC-B+C+D5', en el plásmido pBAC-TcDNA<sup>-ΔClaI</sup> para generar el plásmido pBAC-TcDNA<sup>FL</sup> que contiene el cDNA que codifica para el genoma completo del VGPT (Figura 3).

La estabilidad en bacterias de los plásmidos utilizados en la construcción del clon de cDNA infectivo (pBAC-TcDNA<sup>-ΔClaI</sup> y pBAC-ClaI<sup>F</sup>), así como el plásmido que contiene el cDNA completo (pBAC-TcDNA<sup>FL</sup>) fue analizada después de ser crecidos en *E. coli* durante 160 generaciones. El análisis de estabilidad fue llevado a cabo mediante digestión con enzimas de restricción de los DNAs purificados. No se detectaron deleciones o inserciones, aunque no se puede descartar la presencia de cambios menores no detectados por la técnica de análisis utilizada, en el caso de los plásmidos pBAC-TcDNA<sup>-ΔClaI</sup>, y pBAC-B+C+D5'. En el caso del plásmido pBAC-TcDNA, que contiene el genoma completo de VGPT, se detectó cierta inestabilidad después de 40 generaciones (en este punto aproximadamente el 80 % del DNA presentaba el patrón de restricción correcto). Esta ligera inestabilidad, sin embargo, no representa un obstáculo para el rescate del virus infectivo, dado que 20 generaciones (4 litros de cultivo) de crecimiento bacteriano son suficientes para generar una cantidad de DNA plasmídico que permita rescatar el virus.

### 1.2 Rescate de un VGPT infectivo a partir de un cDNA que codifica para el genoma completo

Células ST fueron transfectadas con el plásmido pBAC -TcDNA<sup>FL</sup>. A las 48 h post-transfección el sobrenadante de cultivo fue recogido y pasado en células ST seis veces (véase la Figura S). A partir del pase 2, a las 14 h post-infección el efecto citopático empezaba a ser aparente para posteriormente, a las 20 h post-infección, hacerse extensible a prácticamente la totalidad de las células que formaban la monocapa (véase la Figura 6). Por otra parte, el título del virus rescatado incremento rápidamente con los pases llegando a valores del orden de 10<sup>8</sup> ufp/ml a partir del pase 3 (véase la Figura 7). El experimento fue repetido cinco veces y en todos los casos se recuperó virus infectivo con títulos similares, mientras que en el caso de células ST no transfectadas o transfectadas con un plásmido similar donde el fragmento ClaI-ClaI se encontraba en la orientación contraria nunca se recuperó virus.

Con el fin de eliminar la posibilidad de que el virus obtenido fuera el producto de una contaminación, la secuencia en las posiciones 6.752 y 18.997 se determinó mediante secuenciación de fragmentos de cDNA amplificados por RT-PCR utilizando como molde el RNA genómico del virus rescatado. El análisis de la secuencia determinó que los nucleótidos en las posiciones 6.752 y 18.997 eran aquéllos presentes en el cDNA. Además, el virus rescatado presentaba en la secuencia del cDNA del gen S un sitio de restricción DraIII en la posición 20.990, como se esperaba para el gen S de C11 (Figura 8). La presencia de estos tres marcadores genéticos confirmaba que el virus aislado procedía del cDNA.

En una caracterización más profunda del virus generado, se llevó a cabo un análisis comparativo por inmunofluorescencia de células infectadas con el virus recuperado (TcDNA) después de la transfección con el plásmido pBAC-TcDNA<sup>FL</sup> o células infectadas con el aislado PUR46-MAD del VGPT. Para ello, se utilizaron anticuerpos policlonales y monoclonales específicos que reconocían tanto el aislado C11 como el PUR46-MAD o solamente este último (véase la Figura 10). Los resultados obtenidos confirmaron la antigenicidad esperada para el nuevo virus TcDNA. El anticuerpo policlonal específico para el VGPT, los monoclonales esperados específicos de la proteína S (ID.B12 y 6A.C3), así como los monoclonales específicos de las proteínas M (3B.B3) y N (3B.D8), reconocían tanto el TcDNA como el PUR46 -MAD. Los datos obtenidos indicaban que el virus generado presentaba las proteínas M y N del aislado PUR46-MAD y la proteína S del aislado C11, como había sido diseñado en el cDNA original.

### 1.3 Infectividad *in vivo* y virulencia

Con el fin de analizar la infectividad *in vivo* del virus TcDNA, un grupo de cinco cerdos recién nacidos fueron inoculados con virus clonado del pase 6, y la mortalidad analizada. Los cinco cerdos inoculados murieron entre los días 3 y 4 post -inoculación indicando que el virus TcDNA era virulento. En contraste dos lechones inoculados solo con el diluyente del virus y mantenidos en las mismas condiciones no sufrieron alteraciones.

#### 1.4 Optimización de los niveles de expresión mediante la modificación de las secuencias reguladoras de la transcripción

La síntesis de RNA en coronavirus tiene lugar mediante un proceso dependiente de RNA, en el cual los mRNAs son transcritos a partir de moldes con polaridad negativa. En el VGPT aparece una secuencia consenso, CTAAACC, la cual se localiza justo por delante de la mayoría de los genes. Estas secuencias representan señales para la transcripción de los mRNAs subgenómicos. En coronavirus existen entre seis u ocho tipos de mRNAs con tamaños variables, dependiendo del tipo de coronavirus y del hospedador. El de mayor tamaño se corresponde con el RNA genómico, el cual a su vez sirve como mRNA para las ORFs 1a y 1b. El resto de mRNAs se corresponden con mRNAs subgenómicos. Estos RNAs se denominan mRNA 1 hasta 7, en orden decreciente de tamaño. Por otra parte algunos mRNAs que han sido descubiertos con posterioridad al conjunto de mRNAs descritos originalmente, han sido denominados con el nombre del mRNA correspondiente, un guión y un número, e.j., mRNA 2-1. Los mRNAs presentan una estructura coterminal en relación con la estructura del RNA genómico. Con la excepción del mRNA de menor tamaño, el resto son estructuralmente policistrónicos, donde en general sólo la ORF localizada más hacia el 5' es traducida.

Se ha estudiado la eficiencia en la expresión de un gen marcador (GUS) utilizando diferentes secuencias flanqueantes al 5' terminal de la secuencia intergénica mínima (IG) CTAAAC (Figura 11), diferentes secuencias flanqueantes al 3' terminal de la secuencia IG (Figura 12), y varios sitios de inserción (Figura 13). Los resultados obtenidos (Figuras 11 a 13) indicaron que la expresión óptima se conseguía con una TRS constituida por: (i) los -88 nt flanqueantes de la secuencia consenso para el gen N del VGPT; (ii) la secuencia IG; y (iii) la secuencia flanqueante al 3' de la secuencia IG del gen S. Además, de acuerdo con los resultados obtenidos según el punto de inserción del gen heterólogo, los mayores niveles de expresión se conseguían cuando el gen heterólogo se localizaba en el extremo 3' del genoma. Una TRS como la descrita permite la expresión de GUS a unos niveles de entre 2 y 8  $\mu\text{g}$  por  $10^6$  células.

#### 1.5 Especificidad de tejido del sistema de expresión

Muchos de los patógenos entran en el hospedador a través de las mucosas. Con el fin de prevenir este tipo de infecciones es importante desarrollar sistemas de expresión que permitan inducir niveles elevados de inmunidad secretora. Esto puede ser conseguido fundamentalmente mediante la administración de los antígenos en los nódulos linfáticos asociados al tracto respiratorio y/o entérico. Para conseguir este fin, y en general para dirigir la expresión de un gen al tejido de interés, se han estudiado las bases moleculares del tropismo del VGPT. Estos estudios han puesto de manifiesto que la especificidad de tejido del VGPT puede ser modificada mediante la construcción de virus recombinantes que contengan el gen S del coronavirus con el tropismo deseado [Ballesteros et al., 1997; Sánchez et al., 1999]. Esta información permite construir sistemas de expresión basados en genomas cDNAs de coronavirus con tropismo respiratorio o entérico.

#### 1.6 Expresión del antígeno viral codificado por la ORF5 del PRRSV utilizando el cDNA infectivo

Con el fin de optimizar los niveles de expresión de genes heterólogos, se hicieron construcciones a partir de un vector de módulos intercambiables flanqueados por secuencias de clonaje que facilitan el intercambio de TRSs y genes heterólogos dentro del vector. La construcción que incluía la ORF 5 del PRRSV flanqueado en su extremo 5' por la secuencia consenso mínima IGS (CUAAAC) precedida de los -88 nts flanqueantes del gen de la nucleocápsida viral (N), y en su extremo 3' por el sitio de restricción SalI (GTCGAC) y una secuencia análoga a la de Kozak (AC)GACC, dieron una expresión óptima (entorno a  $10 \mu\text{g}/10^6$  células). Estos niveles de expresión del gen heterólogo son en principio más que suficientes para la inducción de una respuesta inmune. El gen heterólogo se insertó en la posición previamente ocupada por los genes 3a y 3b del virus, que son dispensables.

#### Depósito de microorganismos

La bacteria derivada de *Escherichia coli*, portadora de un plásmido con el clon infectivo de la invención, identificada como *Escherichia coli* pBAC-TcDNA<sup>FL</sup>, se ha depositado en la Colección Española de Cultivos Tipo (CECT), Burjassot (Valencia), con fecha 24 de noviembre de 1999, correspondiéndole el número de depósito CECT 5265.

#### Bibliografía

Ahlquist, P., R. French, M. Janda, and L.S. Loes-Fries. (1984). "Multicomponent RNA plant virus infection derived from cloned viral cDNA". *Proc. Natl. Acad. Sci. USA*. 81:7066-7070.

## ES 2 170 622 B1

- Ballesteros, M.L., C.M. Sanchez, and L. Enjuanes (1997).** “Two amino acid changes at the N-terminus of transmissible gastroenteritis coronavirus spike protein result in the loss of enteric tropism”. *Virology*. 227:378-388.
- 5 **Baron, M.D., and T. Barrett. (1997).** “Rescue of rinderpest virus from cloned cDNA”. *J. Virol.* 71:1265-1271.
- Boyer, J. C., A. L. and Hanni. (1994).** “Infectious transcripts and cDNA clones of RNA viruses”. *Virology*. 198:415-426.
- 10 **Chang, R. Y., M. A. Hofmann, P. B. Sethna, and D. A. Brian. (1994).** “A cis-acting function for the coronavirus leader in defective interfering RNA replication”. *J. Virol.* 68:8223-8231.
- Collins, P.L., M.G. Hill, E. Camargo, H. Grosfeld, R.M. Chanock, and B.R. Murphy. (1995).**  
15 “Production of infectious human respiratory syncytial virus from cloned dCNA confirms an essential role for the transcription elongation factor from the 5’ proximal open reading frame of the M2 mRNA in gene expression and provides a capability for vaccine development”. *Proc. Natl. Acad. Sci. USA.* 92:11563-11567.
- 20 **Davis, H.L., L.V. Willis, J.F. Smith, and R.E. Johnston. (1989).** “In vitro synthesis of infectious Venezuelan equine encephalitis virus RNA from a cDNA clone: analysis of a viable deletion mutant” . *Virology*. 171:189-204.
- Dubensky, J., T.W., D.A. Driver, J. M. Polo, B.A. Belli, E.M. Latham, C.E. Ibanez, S. Chada,**  
25 **D. Brumm, T.A. Banks, S.J. Mento, D.J. Jolly, and S.M.W. Chang. (1996).** “Sindbis virus DNA-based expression vectors: utility for in vitro and in vivo gene transfer”. *J. Virol.* 70:508-519.
- Durbin, A.P., S.L. Hall, J.W. Siew, S.S. Whitehead, P.L. Collins, and B.R. Murphy. (1997).**  
“Recovery of infectious human parainfluenza virus type 3 from cDNA”. *Virology*. 235:323-332.
- 30 **Enjuanes, L., S.G. Siddell, and W.J. Spaan. 1998.** “Coronaviruses and Arteriviruses”. *Plenum Press*, New York.
- Enjuanes, L., and B.A.M. Van der Zeijst. 1995.** “Molecular basis of transmissible gastroenteritis coronavirus epidemiology”. *In The Coronaviridae*. S.G. Siddell, editor. Plenum Press, New York. 337-376.
- 35 **Frolov, I., T.A. Hoffman, B.N. Prágai, S.A. Dryga, H.V. Huang, S. Schlesinger, and C.M. Rice. (1996).** “Alphavirus-based expression vectors: Strategies and applications”. *Proc. Natl. Acad. Sci. USA.* 93:11371-11377.
- 40 **Garcin, D., T. Pelet, P. Calain, L. Roux, J. Curran, and D. Kolakofsky. (1995).** “A highly recombinogenic system for the recovery of infectious sendai paramyxovirus from cDNA: generation of a novel”. *EMBO J.* 14:6087-6094.
- 45 **Geigenmuller, U., N.B. Ginzton, and S.M. Matsui. (1997).** “Construction of a genome-length cDNA clone for human astrovirus serotype 1 and sintesis of infectious RNA transcripts”. *J. Virol.* 71:1713-1717.
- 50 **Izeta. A., C. Smerdou, S. Alonso, Z. Penzes, A. Mendes, J. Plana-Duran, and L. Enjuanes. (1999).** “Replication and packaging of transmissible gastroenteritis coronavirus-derived synthetic mini-genomes”. *J. Virol.* 73:1535-1545.
- Kim U.-J., H. Shizuya, P. de Jong, B.W. Birren, and M.I. Simon. (1992).** “Stable propagation  
55 of cosmid-sized human DNA inserts in an F-factor based vector”. *Nucleic Acids Res.* 20:1083-1085.
- Lai, C.-J., B. Zhao, E. Hori, and M. Bray. (1991).** “Infectious RNA transcribed from stably cloned full-length cDNA of dengue type 4 virus”. *Proc. Natl. Acad. Sci. USA.* 88:5139-5143.
- 60 **Lai, M.M.C., and D. Cavanagh. (1997).** “The molecular biology of coronaviruses”. *Adv. Virus Res.* 48:1-100.

## ES 2 170 622 B1

- Lai, M.M.C., C.-L. Liao, Y.-J. Lin, and X. Zhang.** (1994). "Coronavirus: how a large RNA viral genome is replicated and transcribed". *Infect. Agents Dis.* 3:98-105.
- Liljeström, P.** (1994). "Alphavirus expression systems". *Curr. Opin. Biotech.* 5:495-500.
- 5 **Liljeström, P., and B. Garoff.** (1991). "A new generation of animal cell expression vectors based on the Semliki Forest virus replicon". *BiolTechnology.* 9:1356-1361.
- Luytjes, W., M. Krystal, M. Enami, J.D. Parvin, and P. Palese.** (1989). "Amplification, expres-  
10 sion, and packaging of a foreign gene by influenza virus". *Cell.* 59:1107-1113.
- Mandl, C.W., M. Ecker, H. Holzmann, C. Kunz, F.X. Heinz.** (1997). "Infectious cDNA clones of tick-borne encephalitis virus European subtype prototypic strain Neudoerfl and high virulence strain Hypr". *J. Gen. Virol.* 78:1049-1057.
- 15 **Maniatis, T., E. P. Fritsh, and J. Sambrook,** (1989). *Molecular cloning: a laboratory manual.* Cold Spring Harbour Laboratory Press. New York.
- Mendez, A., C. Smerdou, A. Izeta, F. Gebauer, and L. Enjuanes.** (1996). "Molecular cha-  
20 racterization of transmissible gastroenteritis coronavirus defective interfering genomes: packaging and heterogeneity". *Virology.* 217:495-507.
- Pushko, P., M. Parker, G.V. Ludwing, N.L. Davis, R.X. Johnston, and J.F. Smith.** (1997).  
25 "Replication-helper systems from attenuated Venezuelan equine encephalitis virus: expression of hetero-  
logous genes in vitro and immunization against heterologous pathogens in vivo". *Virology* 239:389-401.
- Racaniello, V.R., and D. Baltimore.** (1981). "Cloned poliovirus cDNA is infectious in mammalian  
cells". *Science.* 214:916-919.
- 30 **Radecke, F., P. Spielhofer, H. Schneider, K. Kaelin, M. Huber, C. Dotsch, G. Christiansen,  
and M.A. Billeter.** (1995). "Rescue of measles viruses form cloned DNA". *EMBO J.* 14:5773-5784.
- Rice, C.M., A. Grakoui, R. Galler, and T.J. Chambers.** (1989). "Transcription of infectious  
35 yellow fever RNA from full-length cDNA templates produced by in vitro ligation". *New Biologist.* 1:285-  
296.
- Rice, C.M., R. Levis, J.H. Strauss, and H.V. Huang.** (1987). "Production of infectious RNA tran-  
scripts from Sindbis virus cDNA clones: Mapping of lethal mutations, rescue of a temperature-sensitive  
40 marker, and in vitro mutagenesis to generate defined mutants". *J. Virol.* 61:3809-3819.
- Rice, C.M., and J.H. Strauss.** (1981). "Synthesis, cleavage, and sequence analysis of DNA comple-  
mentary to the 26S messenger RNA of Sindbis virus". *J. Mol. Biol.* 150:315-340.
- 45 **Ruggli, N., J.D. Tratschin, C. Mittelholzer, M.A. Hofmann.** (1996). "Nucleotide sequence of  
classical swine fever virus strain Alfort/187 and transcription of infectious RNA from stably cloned full-  
length cDNA". *J. Virol.* 70:3479-3487.
- Sánchez, C. M., G. Jiménez, M. D. Labiada, I. Correa, C. Suñé, M. J. Bullido, F. Gebauer,  
C. Smerdou, P. Callebaut, J. M. Escribano, and L. Enjuanes.** (1990). "Antigenic homology among  
50 coronaviruses related to transmissible gastroenteritis virus". *Virology.* 174:410-417.
- Sánchez, C.M., F. Gebauer, C. Suñé, A. Mendez, J. Dopazo, and L. Enjuanes.** (1992). "Ge-  
netic evolution and tropism of transmissible gastroenteritis coronaviruses". *Virology.* 190:92-105.
- 55 **Sánchez, C.M., A. Izeta, J.M. Sanchez-Morgado, S. Alonso, I. Sola, M. Balasch, J. Plana-  
Durán, and L. Enjuanes.** (1999). "Targeted recombination demonstrates that the spike gene of tran-  
smissible gastroenteritis coronavirus is a determinant of its enteric tropism and virulence". *J. Virol.*  
73:7607-7618.
- 60 **Sawicki, S.G., and D. L. Sawicki.** (1990). "Coronavirus transcription: subgenomic mouse hepatitis  
virus replicative intermediates function in RNA synthesis". *J. Virol.* 64:1050-1056.



## ES 2 170 622 B1

**Schnell, M.J., T. Mebatsion, and K.-K. Conzelmann.** (1994). "Infectious rabies viruses from cloned cDNA". *EMBO J.* 13:4195-4203.

5 **Sethna, P.B., S.-L. Hung, and D.A. Brian.** (1989). "Coronavirus subgenomic minus-strand RNAs and the potential for mRNA replicons". *Proc. Natl. Acad. Sci. USA.* 96:5626-5630.

**Shizuya, E., B. Birren, U.-J. Kim, V. Mancino, T. Slepak, Y. Tachiiri, and M. Simon.** (1992). "Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector". *Proc. Natl. Acad. Sci. USA.* 89:8794-8797.

10 **Siddell, S.G.** 1995. *The Coronaviridae*. Plenum Press, New York. 418 pp.

**Smerdou, C., and P. Liljestrom.** (1999). "Non-viral amplification systems for gene transfer: vectors based on alphaviruses". *Curr. Opin. Mol. Therap.* 1:244-251.

15 **Taniguchi, M., and F.A.P. Miller.** (1978). "Specific suppressive factors produced by hybridomas derived from the fusion of enriched suppressor T cells and A T lymphoma cell line". *J. Exp. Med.* 148:373-382.

20 **van der Most, R.G., and W.J.M. Spaan** 1995. "Coronavirus replication, transcription, and RNA recombination". In *The Coronaviridae*. S.G. Siddell, editor. Plenum Press, New York. 11-31.

**Wang, K., C. Boysen, B. Shizuya, M.I. Simon, and L. Hood.** (1997). "Complete nucleotide sequence of two generations of a bacterial artificial chromosome cloning vector". *BioTechniques.* 23:992-994.

25 **Woo, S.-S., J. Jiang, B.S. Gill, A.R. Paterson, and R.A. Wing.** (1994). "Construction and characterization of a bacterial artificial chromosome library of *Sorghum bicolor*". *Nucleic Acids Res.* 22:4922-4931.

30 **Zhang, X., C.L. Liao, and M.M.C. Lai.** (1994). "Coronavirus leader RNA regulates and initiates subgenomic mRNA transcription both in trans and in cis". *J. Virol.* 68:4738-4746.

35

40

45

50

55

60

**REIVINDICACIONES**

1. Un clon infectivo derivado de un coronavirus que comprende un DNA complementario (cDNA) que codifica el RNA genómico (gRNA) de un coronavirus clonado bajo una secuencia promotora de la transcripción.
- 5
2. Clon infectivo según la reivindicación 1, en el que dicho coronavirus es un aislado del virus de la gastroenteritis porcina transmisible (VGPT).
3. Clon infectivo según la reivindicación 1, en el que dicho promotor es el promotor de expresión inmediatamente temprano (IE) de citomegalovirus (CMV).
- 10
4. Clon infectivo según la reivindicación 1, en el que dicho cDNA está flanqueado en el extremo 3' por una cola poli(A), la ribozima del virus de la hepatitis delta (HDV), y las secuencias de terminación y poliadenilación de la hormona de crecimiento bovina (BGH).
- 15
5. Clon infectivo según la reivindicación 1, en el que dicho cDNA se ha clonado en un cromosoma artificial bacteriano (BAC).
6. Un procedimiento para la obtención de un clon infectivo según cualquiera de las reivindicaciones 1 a 5, que comprende construir el cDNA que codifica el gRNA de un coronavirus y ensamblar los elementos reguladores de la transcripción.
- 20
7. Procedimiento según la reivindicación 6, en el que la construcción del cDNA que codifica el gRNA de un coronavirus comprende:
- 25
- clonar un genoma defectivo interferente derivado de dicho coronavirus bajo un promotor de expresión en un BAC;
- completar las deleciones de dicho genoma defectivo interferente y regenerar las secuencias delecionadas con respecto al gRNA infectivo;
- 30
- identificar las secuencias tóxicas para la bacteria en la que se va a clonar,
- retirar dichas secuencias tóxicas, e
- 35
- insertar dichas secuencias tóxicas justo antes de efectuar la transfección en células eucarióticas para obtener el clon de cDNA que codifica el gRNA del coronavirus.
8. Un vector viral recombinante que comprende un clon infectivo según cualquiera de las reivindicaciones 1 a 5, u obtenible según el procedimiento de cualquiera de las reivindicaciones 6 ó 7, modificado para contener un ácido nucleico heterólogo insertado en dicho clon infectivo bajo condiciones que permiten la expresión de dicho ácido nucleico heterólogo.
- 40
9. Vector según la reivindicación 8, en el que dicho ácido nucleico heterólogo se selecciona entre un gen y un fragmento de un gen que codifica un producto génico de interés.
- 45
10. Un método para producir un producto de interés que comprende cultivar una célula hospedadora que contiene un vector viral según cualquiera de las reivindicaciones 8 ó 9 bajo condiciones que permiten la expresión del ácido nucleico heterólogo y recuperar el producto de interés.
- 50
11. Un método para producir un coronavirus recombinante modificado, que contiene un ácido nucleico heterólogo en una secuencia de cDNA, correspondiente al genoma de un coronavirus, que comprende introducir un vector viral según cualquiera de las reivindicaciones 8 ó 9 en una célula hospedadora, cultivar dicha célula hospedadora que contiene dicho vector viral bajo condiciones que permiten la expresión y replicación del vector viral y recuperar los viriones obtenidos del coronavirus recombinante modificado.
- 55
12. Una vacuna capaz de inducir protección en un animal frente a la infección causada por un agente infeccioso que comprende (i) al menos, un vector viral según la reivindicación 8, que expresa, al menos, un antígeno adecuado para inducir una respuesta inmune frente a dicho agente infeccioso, o un anticuerpo que proporciona protección contra dicho agente infeccioso, junto con, opcionalmente, (ii) un excipiente farmacéuticamente aceptable.
- 60

## ES 2 170 622 B1

13. Vacuna según la reivindicación 12, en la que dicho vector viral expresa, al menos, un antígeno capaz de inducir una respuesta inmune sistémica y/o una respuesta inmune en mucosas frente a distintos agentes infecciosos que se propagan en mucosas respiratorias o entéricas.

5 14. Una vacuna multivalente capaz de inducir protección en un animal frente a la infección causada por más de un agente infeccioso que comprende (i) un vector viral según la reivindicación 8, que expresa, unos antígenos adecuados para inducir una respuesta inmune frente a dichos agentes infecciosos, o anticuerpos que proporcionan protección contra dichos agentes infecciosos, junto con, opcionalmente, (ii) un excipiente farmacéuticamente aceptable.

10 15. Una vacuna multivalente capaz de inducir protección en un animal frente a la infección causada por más de un agente infeccioso que comprende (i) más de un vector viral según la reivindicación 8, cada uno de los cuales expresa un antígeno adecuado para inducir una respuesta inmune frente a cada uno de dichos agentes infecciosos, o anticuerpos que proporcionan protección contra cada uno de dichos agentes  
15 infecciosos, junto con, opcionalmente, (ii) un excipiente farmacéuticamente aceptable.

16. Un método para producir un coronavirus recombinante que comprende introducir un clon infeccioso según cualquiera de las reivindicaciones 1 a 5, u obtenible según el procedimiento de cualquiera de las reivindicaciones 6 ó 7, en una célula hospedadora, cultivar dicha célula hospedadora que contiene el  
20 clon infeccioso bajo condiciones que permiten la expresión y replicación del clon infeccioso y recuperar los viriones obtenidos del coronavirus recombinante que contienen el genoma completo del coronavirus.

25

30

35

40

45

50

55

60

FIGURA 1

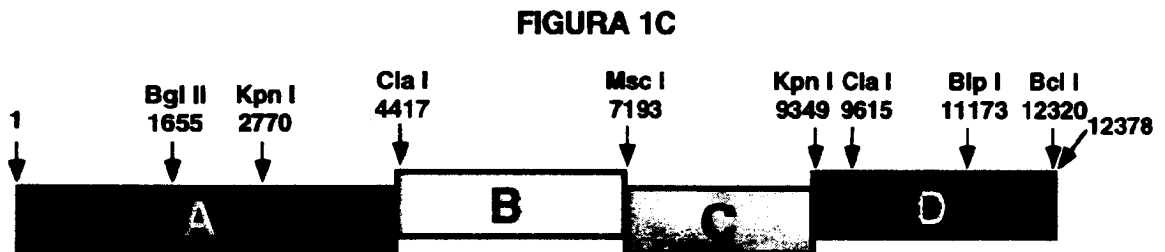
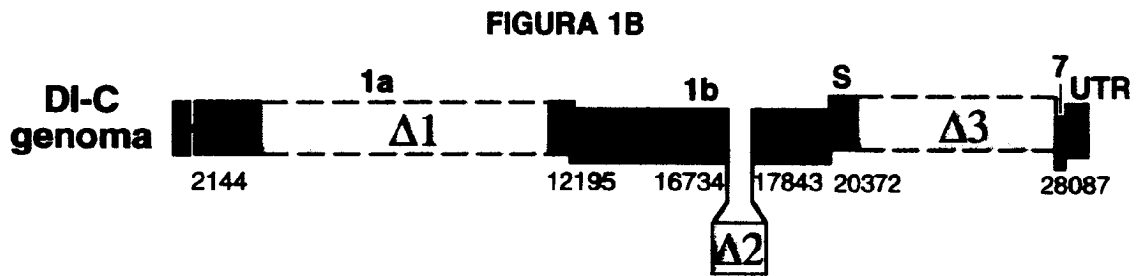
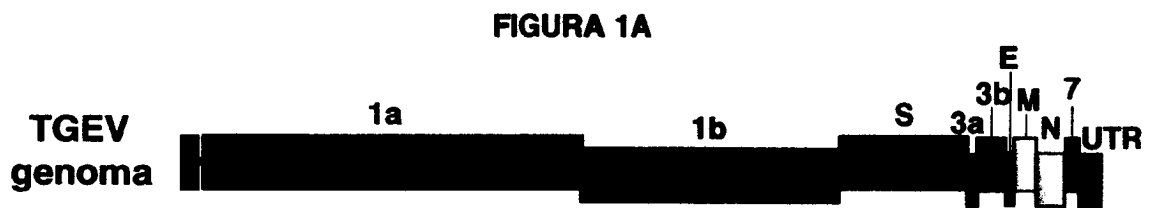


FIGURA 2

# pBeloBAC

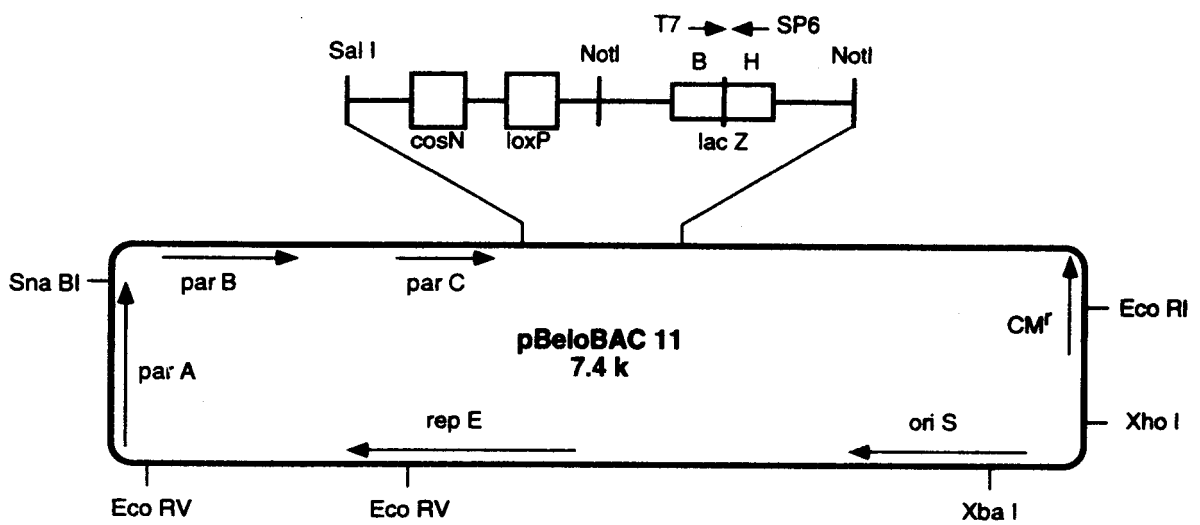
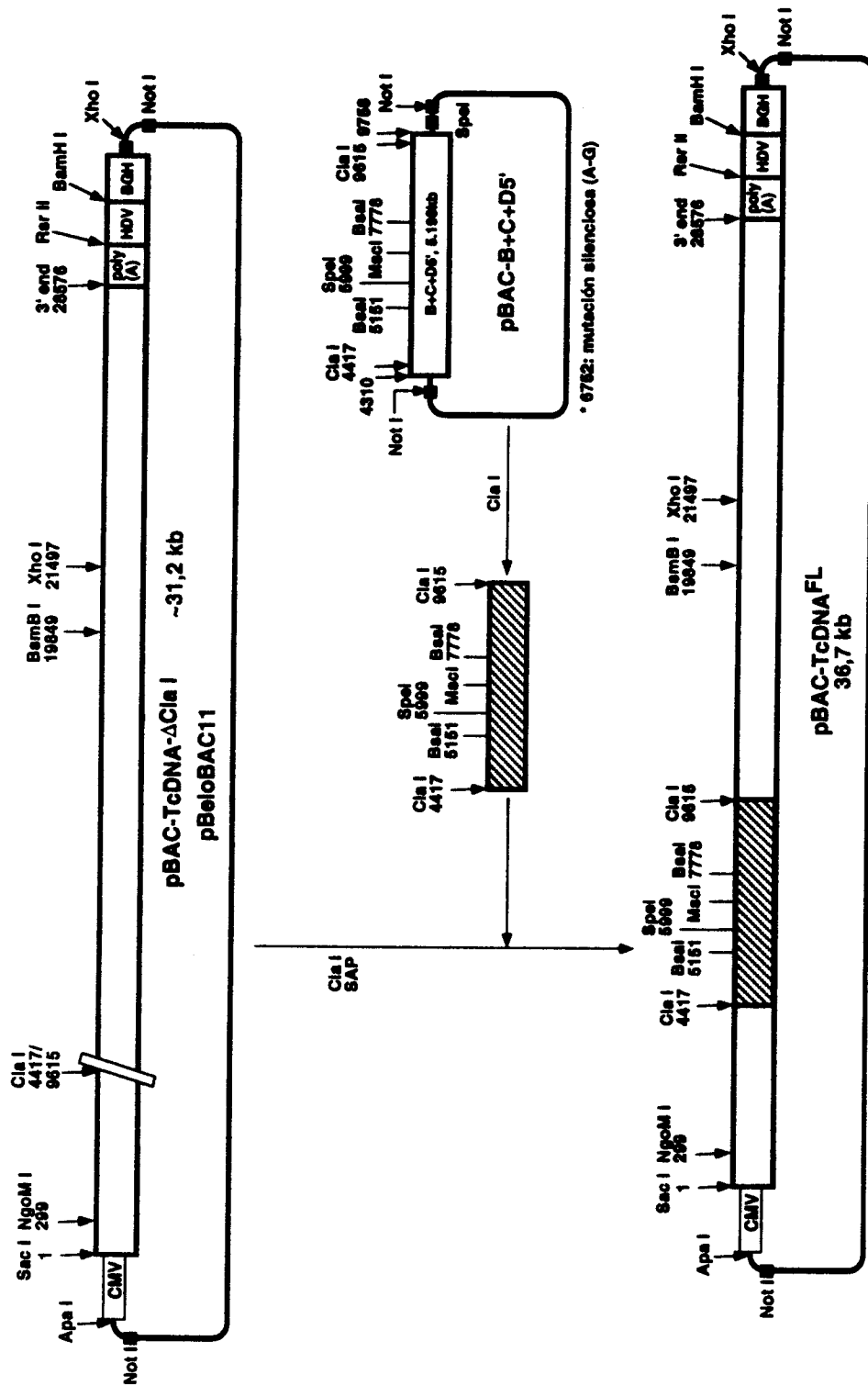


FIGURA 3



**FIGURA 4**

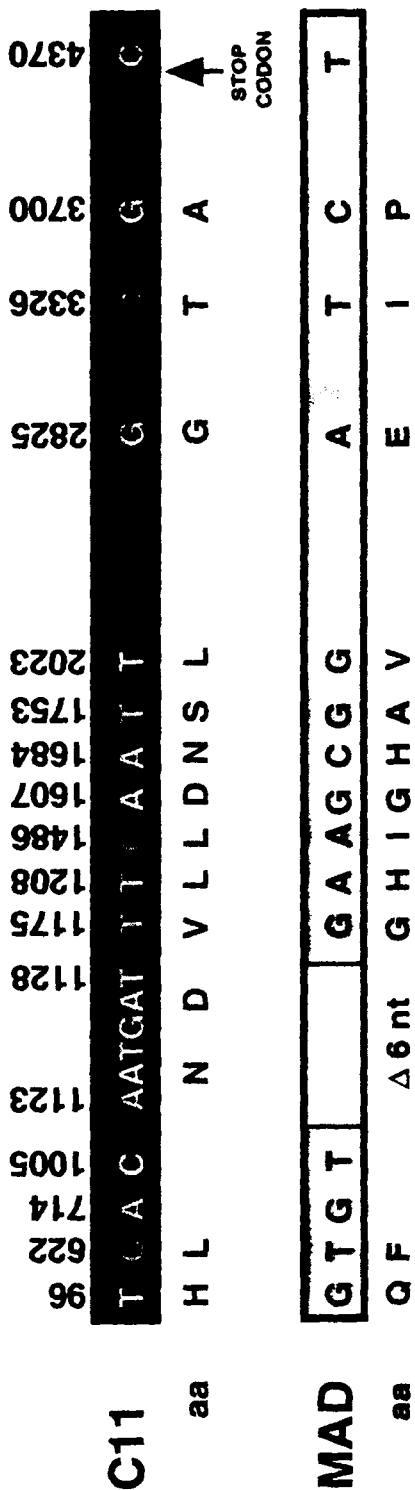
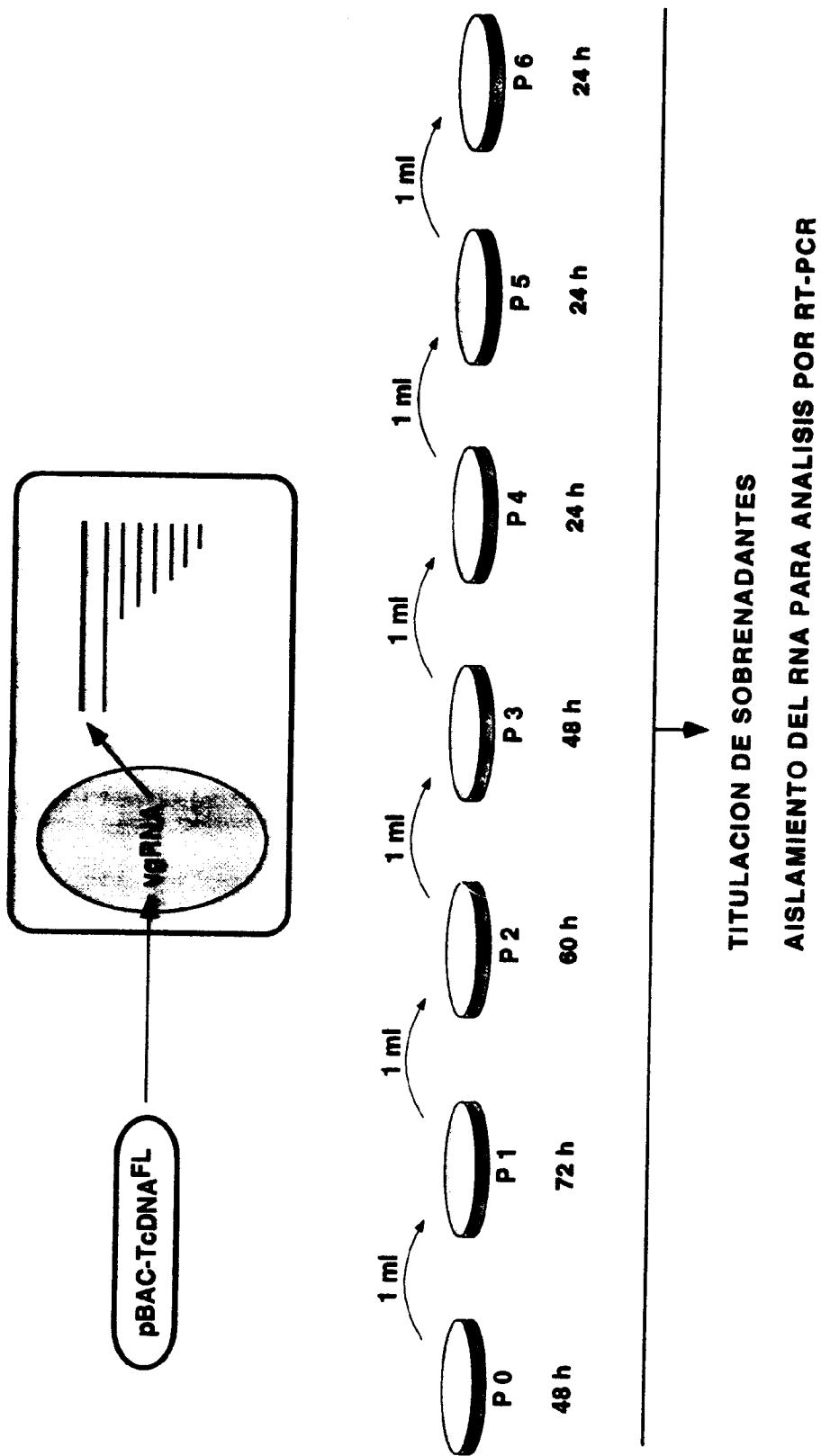


FIGURA 5



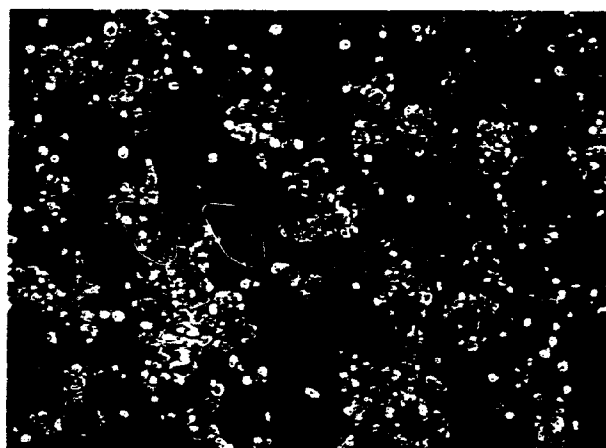


## Figura 6

**Figura 6A**



**Figura 6B**



**Figura 6C**

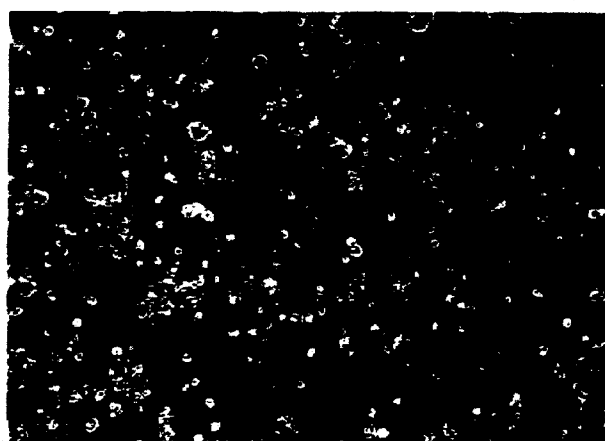


FIGURA 7

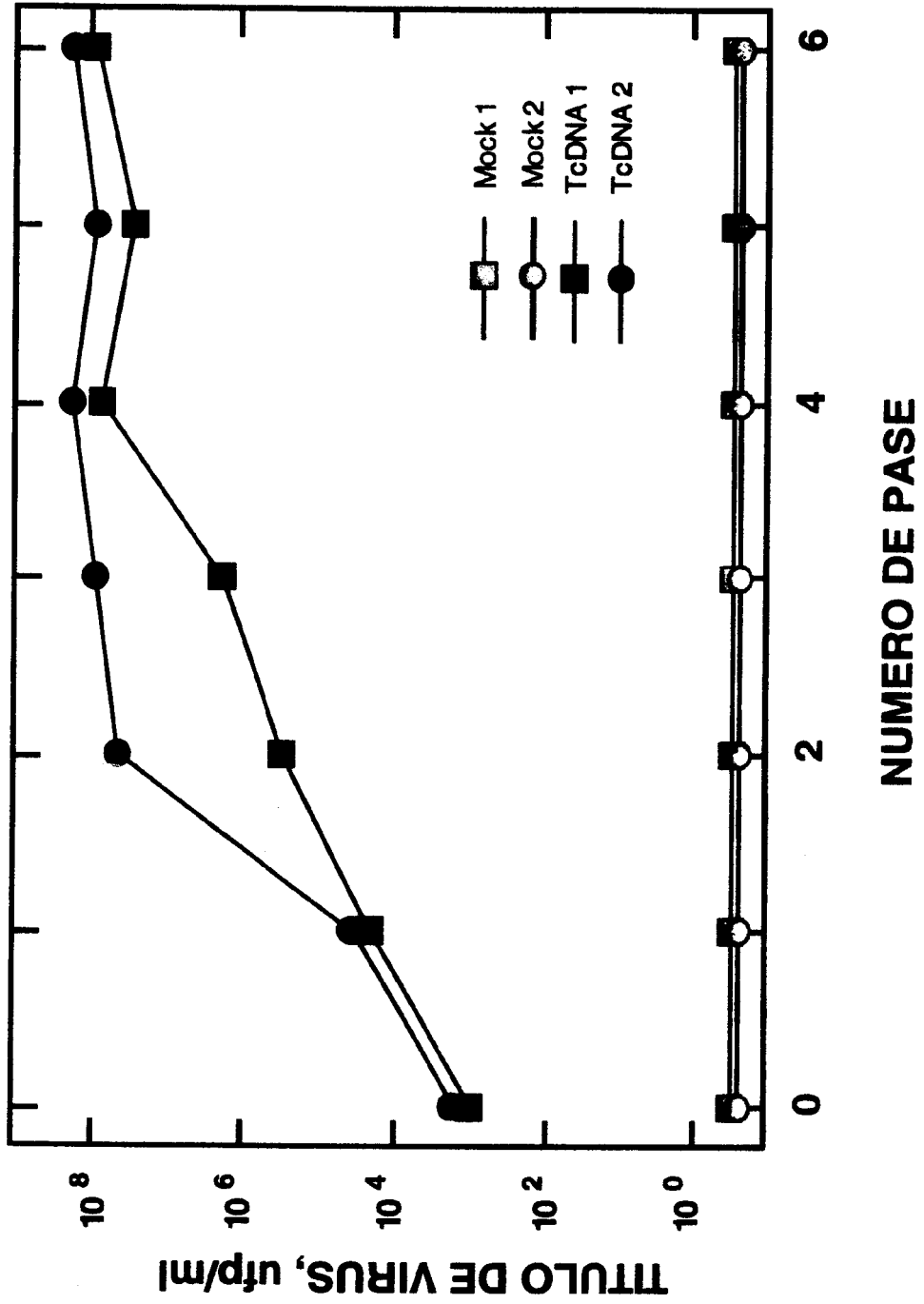
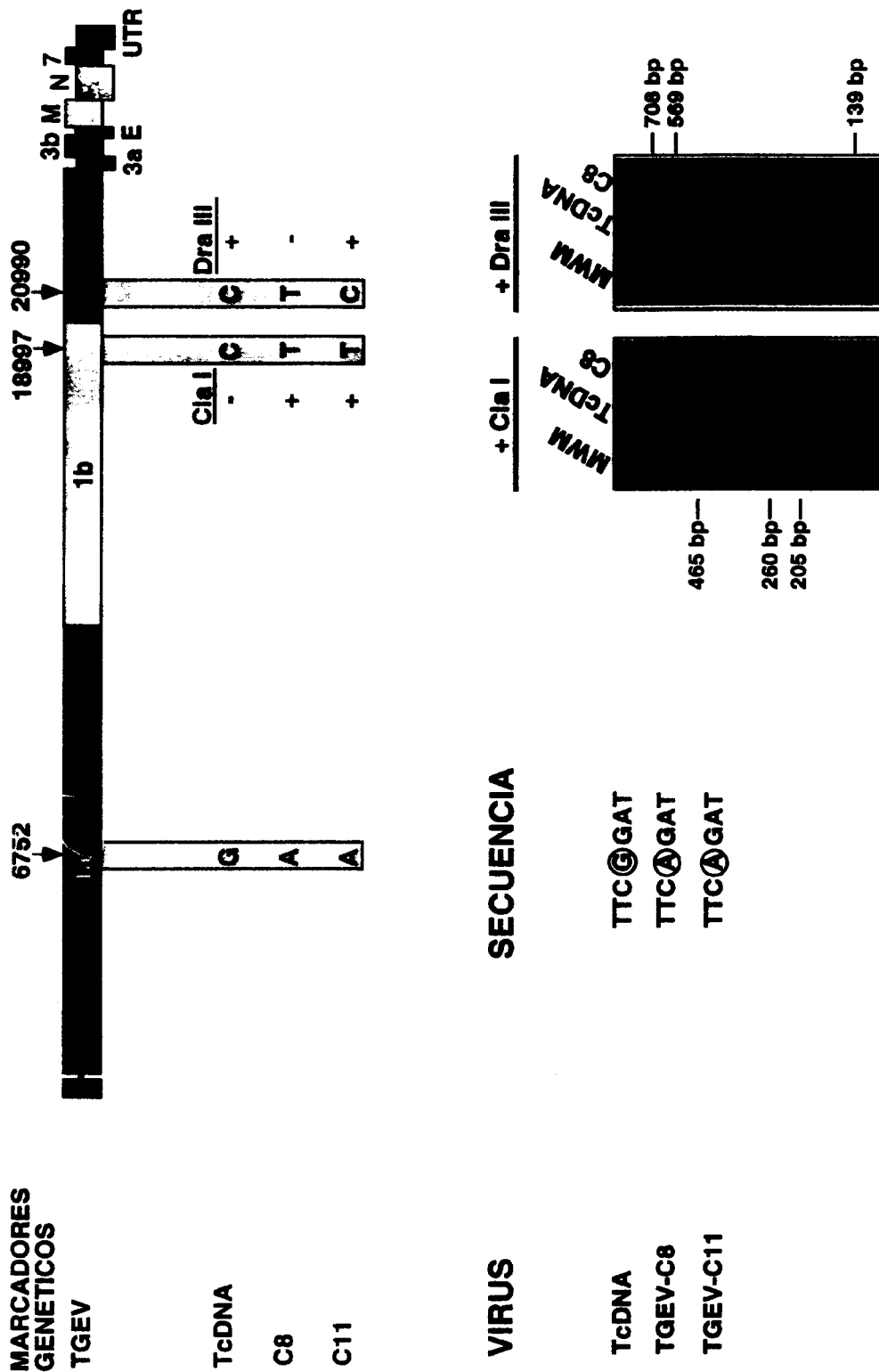
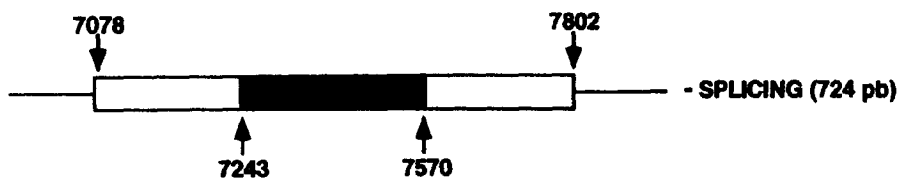


FIGURA 8

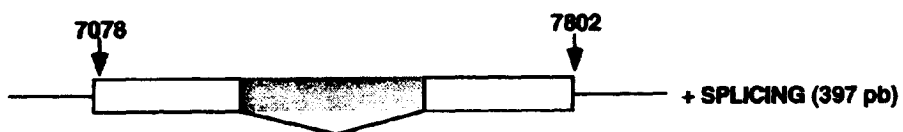


**FIGURA 9**

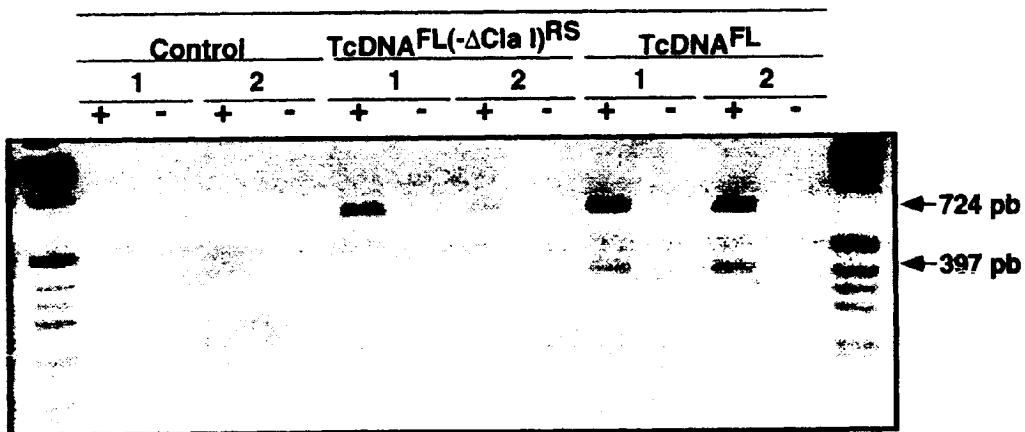
**FIGURA 9A**



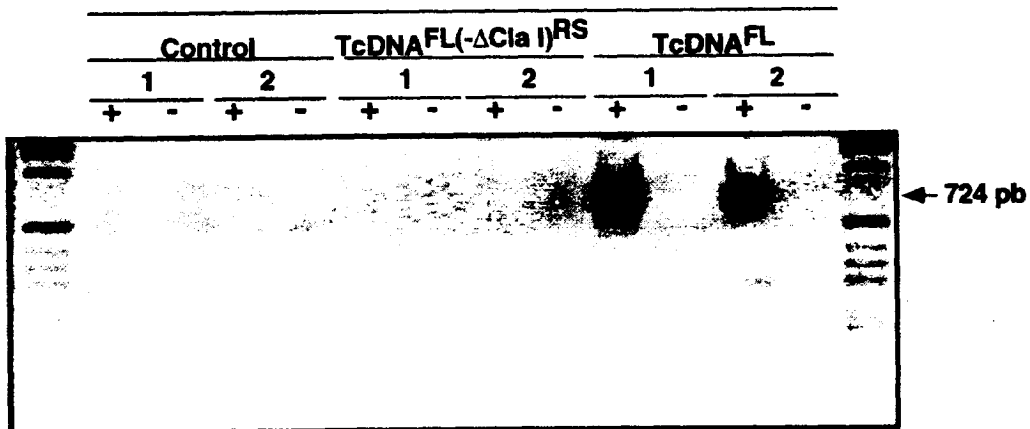
**FIGURA 9B**

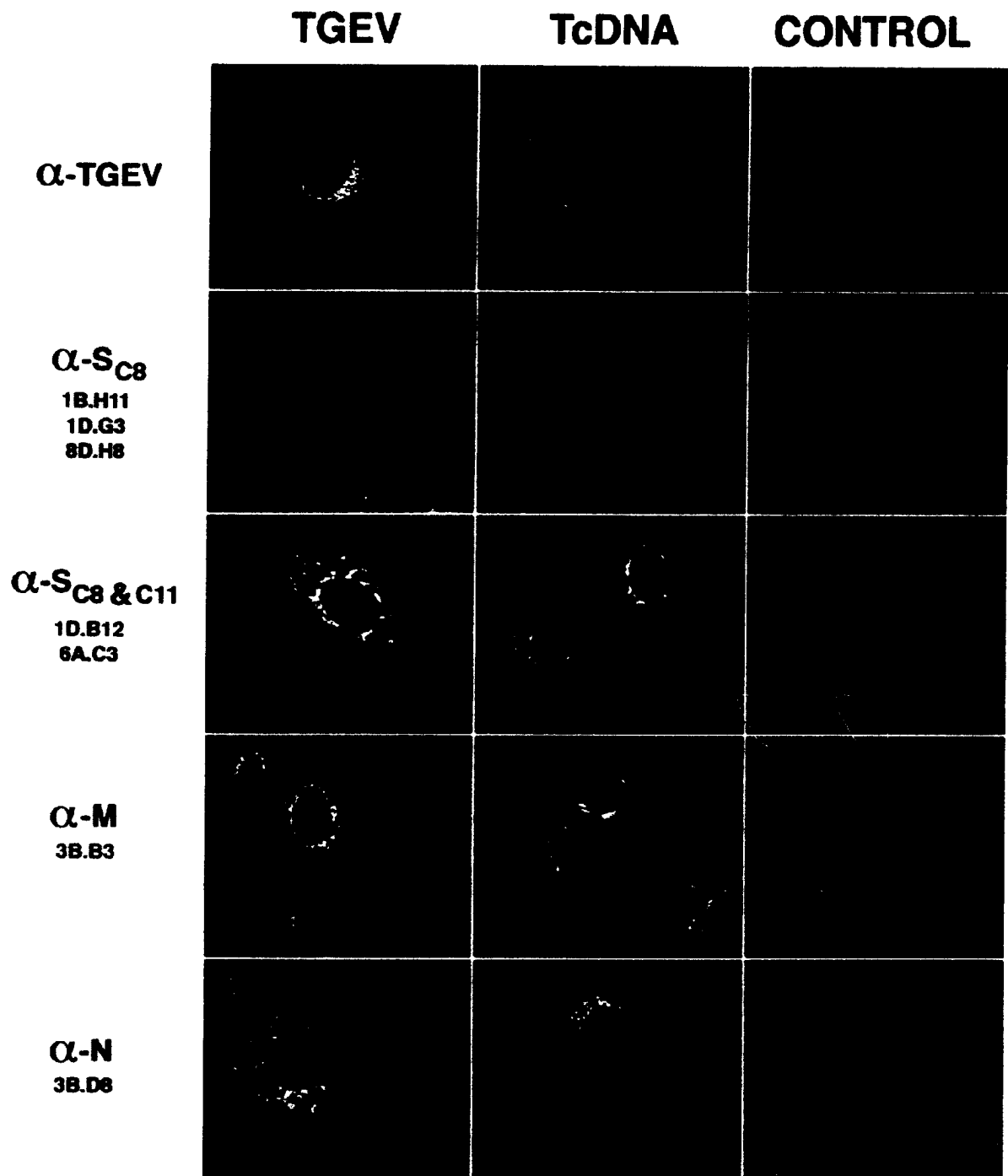


**FIGURA 9C**



**FIGURA 9D**





**Figura 10**

FIGURA 11

FIGURA 11A

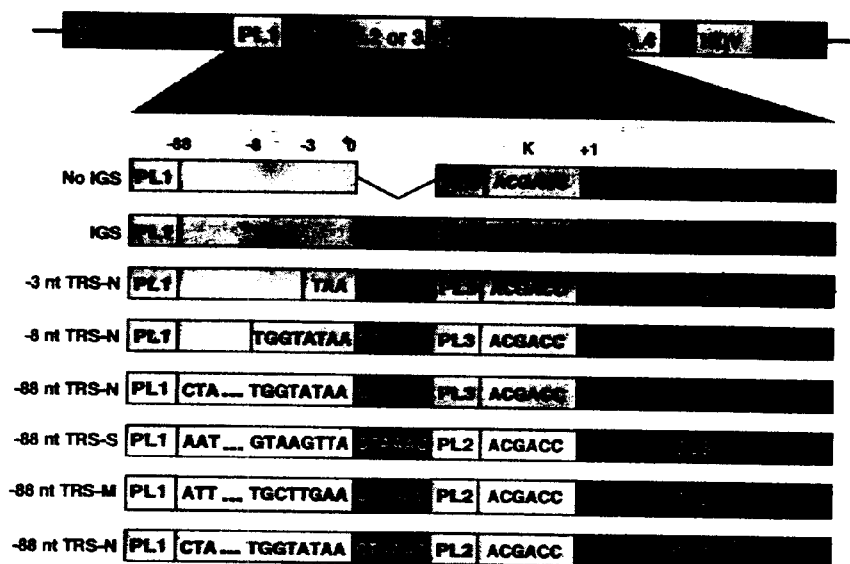


FIGURA 11B

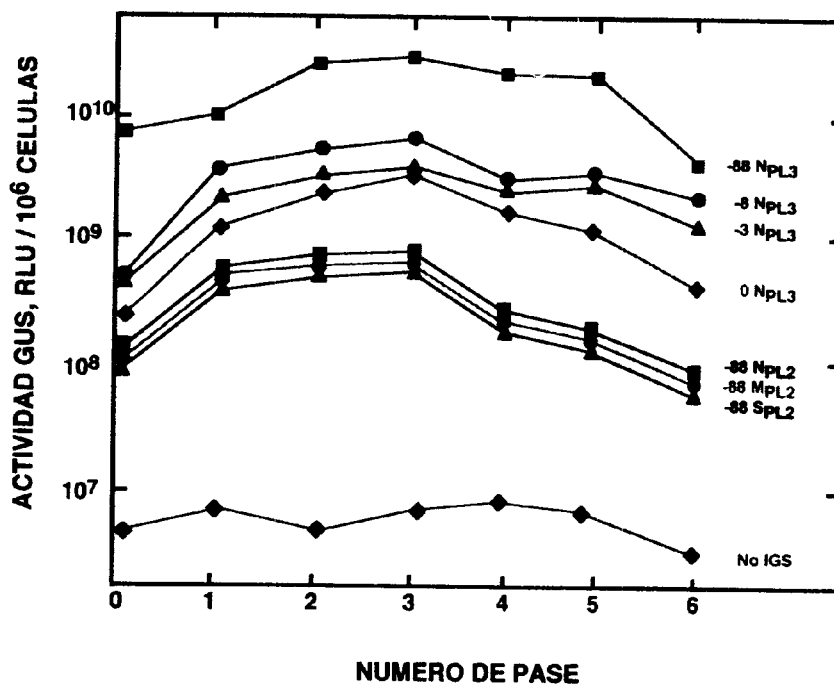


FIGURA 12

FIGURA 12A

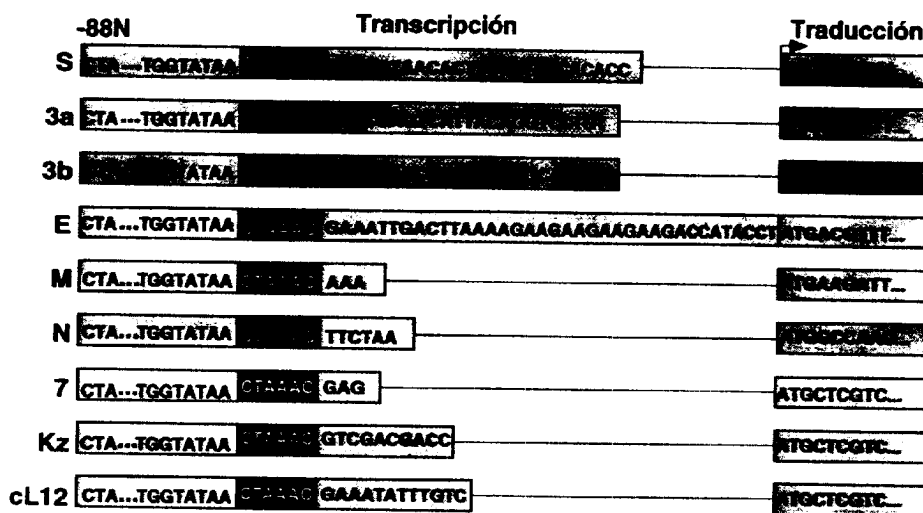
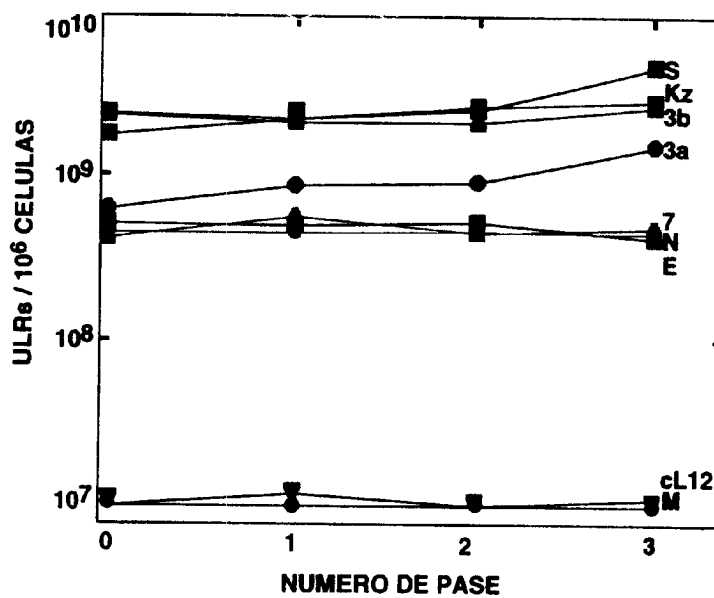
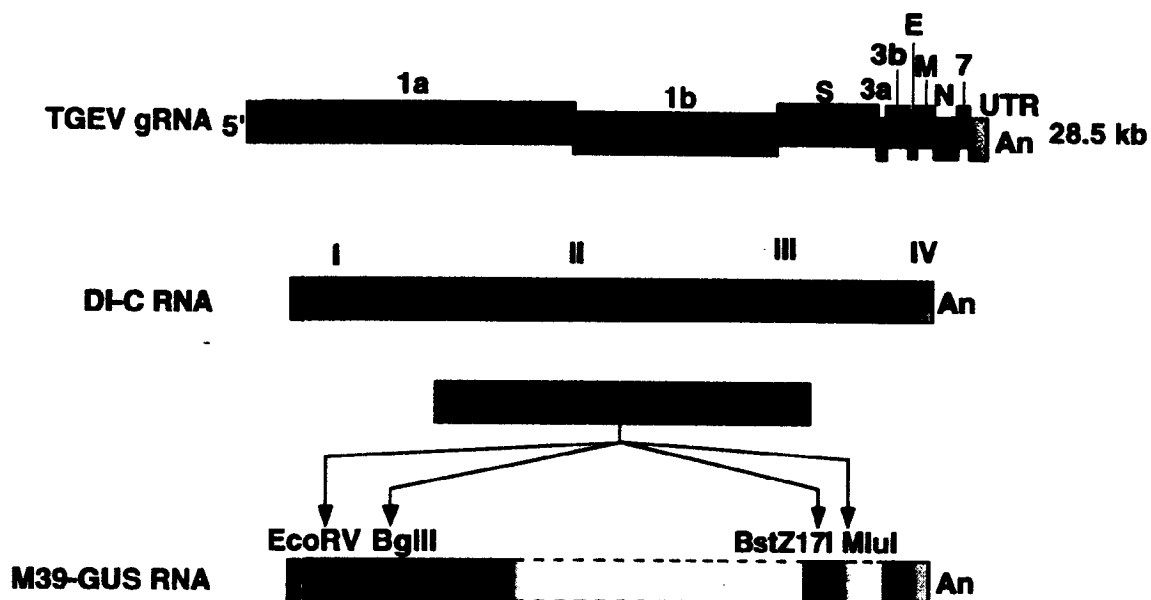


FIGURA 12B

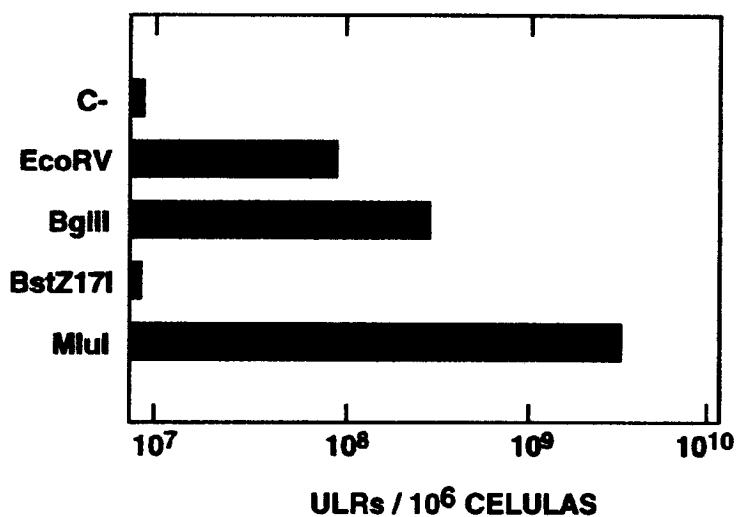


**FIGURA 13**

**FIGURA 13A**



**FIGURA 13B**





# ES 2 170 622 B1

## LISTAS DE SECUENCIAS

### (1) INFORMACION GENERAL:

5 (i) SOLICITANTE:

(A) NOMBRE: CONSEJO SUPERIOR DE INVESTIGACIONES CIENTIFICAS

(B) CALLE: Serrano, 113

10 (C) CIUDAD: Madrid

(D) ESTADO: Madrid

15 (E) PAIS: ES

(F) CODIGO POSTAL (ZIP): 28006

20 (G) TELEFONO: 91 585 50 00

(H) FAX: 91 411 30 77

(ii) TITULO DE LA INVENCION: CLONES Y VECTORES INFECTIVOS DERIVADOS DE CO-  
RONAVIRUS Y SUS APLICACIONES

25 (iii) NUMERO DE SECUENCIAS: 5

(iv) FORMA LEGIBLE POR ORDENADOR:

30 (A) TIPO DE SOPORTE: Floppy disk

(B) ORDENADOR: IBM PC compatible

(C) SISTEMA OPERATIVO: PC-DOS/MS-DOS

35 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (OEP)

### (2) INFORMACION DE LA SEC. ID. N°: 1:

40 (i) CARACTERISTICAS DE LA SECUENCIA:

(A) LONGITUD: 28588 pares de bases

(B) TIPO: ácido nucleico

45 (C) NUMERO DE CADENAS: sencilla

(D) TOPOLOGIA: lineal

50 (ii) TIPO DE MOLECULA: ADNc

(vi) ORIGEN:

(A) ORGANISMO: Virus de la gastroenteritis porcina transmisible

55 (C) AISLADO INDIVIDUAL: PUR46-MAD

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. N°: 1:

60

# ES 2 170 622 B1

	ACTTTTAAAG TAAAGTGAGT GTAGCGTGGC TATATCTCTT CTTTTACTTT AACTAGCCTT	60
5	GTGCTAGATT TTGTCTTCGG ACACCAACTC GAACTAAACG AAATATTTGT CTTTCTATGA	120
	AATCATAGAG GACAAGCGTT GATTATTTCC ATTCAGTTTG GCAATCACTC CTTGGAACGG	180
	GGTTGAGCGA ACGGTGCAGT AGGGTTCCGT CCCTATTTTCG TAAGTCGCCT AGTAGTAGCG	240
10	AGTGCGGTTC CGCCCGTACA ACGTTGGGTA GACCGGGTTC CGTCCTGTGA TCTCCCTCGC	300
	CGGCCGCCAG GAGAATGAGT TCCAAACAAT TCAAGATCCT TGTTAATGAG GACTATCAAG	360
	TCAACGTGCC TAGTCTTCCT ATTCGTGACG TGTTACAGGA AATTAAGTAC TGCTACCGTA	420
	ATGGATTTGA GGGCTATGTT TTCGTACCAG AATACTGTCTG TGACCTAGTT GATTGCGATC	480
15	GTAAGGATCA CTACGTCATT GGTGTTCTTG GTAACGGAGT AAGTGATCTT AAACCTGTTC	540
	TTCTTACCGA ACCCTCCGTC ATGTTGCAAG GCTTTATTGT TAGAGCTAAC TGCAATGGCG	600
	TTCTTGAGGA CTTTGACCTT AAAATTGCTC GCACTGGCAG AGGTGCCATA TATGTTGATC	660
20	AATACATGTG TGGTGCTGAT GGAAAACCAG TCATTGAAGG CGATTTTAAG GACTACTTCG	720
	GTGATGAAGA CATCATTGAA TTTGAAGGAG AGGAGTACCA TTGCGCTTGG ACAACTGTGC	780
	GCGATGAGAA ACCGCTGAAT CAGCAAACCTC TCTTTACCAT TCAGGAAATC CAATACAATC	840
	TGGACATTCC TCATAAATTG CCAAACCTGTG CTACTIONAGACA TGTAGCACCA CCAGTCAAAA	900
25	AGAACTCTAA AATAGTTCTG TCTGAAGATT ACAAGAAGCT TTATGATATC TTCGGATCAC	960
	CCTTTATGGG AAATGGTGAC TGTCTTAGCA AATGCTTTGA CACTCTTCAT TTTATCGCTG	1020
	CTACTCTTAG ATGCCCGTGT GGTCTGAAA GTAGCGGCGT TGGAGATTGG ACTGGTTTTA	1080
30	AGACTGCCGT TTGTGGTCTT TCTGGCAAAG TTAAGGGTGT CACTTTGGGT GATATTAAGC	1140
	CTGGTGATGC TGTGTCACT AGTATGAGCG CAGGTAAGGG AGTTAAGTTC TTTGCCAATT	1200
	GTGTTCTTCA ATATGCTGGT GATGTTGAAG GTGTCTCCAT CTGGAAAGTT ATTAACAACTT	1260
35	TTACAGTTGA TGAGACTGTA TGCACCCCTG GTTTTGAAGG CGAATTGAAC GACTTCATCA	1320
	AACCTGAGAG CAAATCACTA GTTGTCATGCA GCGTTAAAAG AGCATTCACT ACTGGTGATA	1380
	TTGATGATGC TGTACATGAT TGTATCATT CAGGAAAATT GGATCTTAGT ACCAACCTTT	1440
	TTGGTAATGT TGGTCTATTA TTCAAGAAGA CTCCATGGTT TGTACAAAAG TGTGGTGCAC	1500
40	TTTTTGTAGA CGCTTGGAAA GTAGTAGAGG AGCTTTGTGG TTCACTCACA CTTACATAACA	1560
	AGCAAATTTA TGAAGTTGTA GCATCACTTT GCACTTCTGC TTTTACGATT GTAAACTACA	1620
	AGCCAACATT TGTGGTTCCA GACAATCGTG TTAAGATCT TGTAGACAAG TGTGTGAAAAG	1680
45	TTCTTGTAAG AGCATTGAT GTTTTTACGC AGATTATCAC AATAGCTGGT ATTGAGGCCA	1740
	AATGCTTTGT GCTTGGTGCT AAATACCTGT TGTTCAATAA TGCACTTGTC AACTTGTC	1800
	GTGTTAAAAT CCTTGGCAAG AAGCAAAAGG GTCTTGAATG TGCATTCTTT GCTACTAGCT	1860
50	TGGTTGGTGC AACTGTTAAT GTGACACCTA AAAGAACAGA GACTGCCACT ATCAGCTTGA	1920
	ACAAGGTTGA TGATGTTGTA GCACCAGGAG AGGGTTATAT CGTCATTGTT GGTGATATGG	1980
	CTTCTACAA GAGTGGTGAA TATTATTTCA TGATGTCTAG TCCTAATTTT GTTCTTACTA	2040
	ACAATGTTTT TAAAGCAGTT AAAGTTCCAT CTTATGACAT CGTTTATGAT GTTGATAATG	2100
55	ATACCAAAG CAAAATGATT GCAAACCTTG GTTCATCATT TGAATATGAT GGTGATATTG	2160
	ATGCTGCTAT TGTAAGGTC AATGAACTAC TCATTGAATT TAGGCAGCAA AGCTTGTGCT	2220
	TCAGAGCTTT TAAGGACGAC : AAAAGCATTT : TFGTTGAAGC GTATTTTAAA AAGTATAAAA	2280

60

# ES 2 170 622 B1

	TGCCAGCATG	CCTTGCAAAA	CATATTGGTT	TGTGGAACAT	CATAAAGAAA	GATTCATGTA	2340
5	AGAGGGGTTT	TCTTAATCTC	TTCAATCACT	TGAATGAATT	GGAAGATATC	AAAGAAACTA	2400
	ATATTCAGGC	TATTA AAAAC	ATTCTTTGCC	CTGATCCTCT	TCTTGATCTG	GATTATGGTG	2460
	CCATTTGGTA	CAATTGCATG	CCAGGTTGCT	CTGATCCTTC	AGTTTTGGGG	AGTGTCAAC	2520
	TTTTGATCGG	TAATGGTGTG	AAAGTAGTTT	GTGATGGCTG	CAAAGGTTTT	GCTAACCAAC	2580
10	TTTCAAAAGG	TTACAACAAG	CTCTGTAATG	CGGCTCGCAA	TGATATTGAG	ATCGGTGGTA	2640
	TACCATTTTC	CACTTTTAAA	ACACCTACAA	ATACTTTTAT	TGAAATGACA	GATGCTATCT	2700
	ATTCAGTTAT	TGAACAAGGT	AAGGCATTAT	CCTTTAGAGA	TGCTGATGTG	CCAGTTGTAG	2760
	ACAATGGTAC	CATTCTACT	GCTGATTGGT	CTGAACCCAT	TCTGCTTGAA	CCTGCTGAAT	2820
15	ATGTAAAACC	AAAGAACAAT	GGTAATGTCA	TTGTTATTGC	AGGTTATACA	TTTTATAAAG	2880
	ATGAGGATGA	ACATTTTTAT	CCTTATGGTT	TTGGTAAAAT	TGTGCAGAGA	ATGTATAATA	2940
	AAATGGGTGG	TGGTGACAAA	ACTGTCTCAT	TTTCAGAAGA	AGTAGATGTT	CAAGAAATTG	3000
20	CACCTGTTAC	ACGTGTTAAA	CTTGAATTCG	AATTTGACAA	TGAAATTGTA	ACTGGTGTTT	3060
	TTGAACGGGC	TATTGGTACT	AGATACAAAAT	TTACTGGTAC	AACTTGGGAA	GAATTTGAAG	3120
	AGTCTATTTT	TGAAGAAGTC	GATGCAATCT	TTGATACTCT	AGCAAACCAA	GGTGTGCAAC	3180
	TTGAAGGTTA	CTTCATTTAT	GACACTGTGT	GTGGCTTTGA	TATAAAAAAT	CCAGATGGTA	3240
	TTATGATCTC	TCAGTATGAT	ATCAATATTA	CTGCTGATGA	AAAATCAGAA	GTTAGTGCAT	3300
25	CAAGTGAAGA	AGAAGAAGTT	GAATCTGTTG	AAGAAGATCC	TGAGAATGAA	ATTGTAGAAG	3360
	CATCTGAAGG	TGCTGAAGGG	ACTTCTTCTC	AAGAAGAGGT	TGAAACAGTA	GAAGTTGCAG	3420
	ATATTACTTC	TACAGAAGAA	GATGTTGACA	TTGTTGAAGT	ATCTGCTAAA	GATGACCCTT	3480
30	GGGCTGCAGC	TGTTGATGTA	CAAGAAGCTG	AACAATTTAA	TCCTTCTCTA	CCACCTTTCA	3540
	AGACAACGAA	TCTCAACGGA	AAAATTATCC	TTAAGCAAGG	GGATAATAAT	TGTTGGATAA	3600
	ATGCTTGTTG	CTATCAGCTT	CAGGCCTTTG	ATTTTTTCAA	CAATGAAGCT	TGGGAGAAAT	3660
	TTAAGAAAGG	TGATGTCATG	GACTTTGTAA	ACCTTTGTTA	TGCAGCAACA	ACACTAGCAA	3720
35	GAGGTCATTC	TGGTGATGCA	GAGTATCTTC	TTGAACTTAT	GCTCAATGAT	TATAGCACAG	3780
	CCAAGATAGT	ACTTGCAGCT	AAGTGTGGTT	GTGGTGAAAA	AGAAATTGTT	TTGGAAAGAG	3840
	CTGTTTTTAA	ACTCACCCCA	CTTAAGGAGA	GTTTTAATTA	TGGTGTGTTG	GGTGAAGTGA	3900
	TGCAAGTTAA	CACCTGTAGA	TTTTTAAGTG	TTGAAGGCTC	TGGTGTGTTT	GTTTCATGACA	3960
40	TATTAAGCAA	GCAAACGCCA	GAAGCTATGT	TTGTTGTCAA	ACCTGTTATG	CATGCAGTTT	4020
	ACACTGGCAC	AACTCAAAAAT	GGCCATTACA	TGGTTGATGA	TATTGAAACAC	GGTTATTGTG	4080
	TAGATGGTAT	GGGTATTA AAA	CCACTTAAGA	AACGGTGTTA	TACATCCACA	TTGTTTATTAA	4140
	ATGCCAATGT	AATGACTAGA	GCTGAAAAAC	CAAAACAAGA	GTTTAAAGTT	GAAAAAGTAG	4200
45	AACAGCAACC	GATAGTGGAG	GAAAACAAAT	CCTCTATTGA	AAAAGAGGAA	ATTCAAAGTC	4260
	CTAAAAACGA	TGACCTTATA	CTTCCATTTT	ACAAAGCTGG	TAAACTTTCC	TTTTATCAGG	4320
	GTGCTTTTGA	TGTTTTGATC	AATTTCTTGG	AACCTGATGT	TATTGTTAAT	GCTGCTAATG	4380
	GTGATCTTAA	ACACATGGGT	GGTGTGCGAA	GAGCCATCGA	TGTTTTCACT	GGTGGCAAAT	4440
50	TAACAGAACG	TTCTAAGGAT	TATCTTAAAA	AGAACAAATC	TATTGCTCCT	GGTAATGCTG	4500
	TTTTCTTTGA	AAATGTCATT	GAGCATCTTA	GTGTTTTGAA	TGCAGTTGGA	CCACGTAATG	4560
	GTGACAGCCG	AGTTGAAGCC	AACTTTTGTA	ATGTTTACAA	AGCAATTGCA	AAGTGTGAAG	4620
	GAAAAATATT	AACACCACTT	ATTAGTGTGT	GTATCTTTAA	TGTTAGACTT	GAAACATCAT	4680
55	TGCAGTGCTT	ACTTAAGACT	GTGAATGACA	GGGGATTGAA	TGTCTTCGTA	TACACTGACC	4740
	AGGAGAGGCA	AACTATTGAG	AATTTCTTCT	CTTGTCTTAT	CCCTGTCAAT	GTTACTGAGG	4800
	ATAATGTTAA	CGATGACGCT	GTGTCTGTTT	CTTTTGACAA	AACATACGGT	GAAACAGCTTA	4860

60

# ES 2 170 622 B1

	AGGGCACC	GTGTCAT	CACAAAG	ATGTTAC	AAACCA	GTTGCCT	AGC	GCTTTTG	ATG	4920	
5	TTGGTCAA	AAAG	AGTTAT	TAAAG	GCTATT	GATA	TAGATT	TGGCA	AGCTCATT	TAT	4980
	ATGCTGCT	GC	TTTTAG	CGCT	AGTAGT	CATG	ATGCTT	TATAA	ATTTGA	AAGTT	5040
	GCAATTT	CAT	TGTGCAT	AAG	CAGACT	GACA	ACAAC	TGTTG	GATTAAT	GCA	5100
	CATTACAG	AG	ACTCAAG	CCA	CAGTGG	AAAT	TTCCTG	TGT	TAGAGG	TCTC	5160
10	TTCTTGAG	CG	TAAACACA	AA	GGTTTT	GTAC	ATATGT	TGTA	TCACATT	TCT	5220
	AAGGTGAG	CC	AGGTGAT	GCT	GAATTA	ATGC	TGCATA	AACT	TGGTGAC	TG	5280
	ATTGTGAA	AT	CATTGTC	ACA	CACACT	TACAG	CATGTG	ACAA	GTGCGC	AAAA	5340
	TTGTTGG	ACC	AGTGGT	AGCA	GCACCT	CCTTG	CAATTC	ATG	CACTGAC	GAA	5400
15	ATGGCGTT	AG	TGTCAAT	GTG	AAAGT	CACCC	AAATTA	AGGG	CACTGTT	GCT	5460
	TGATTGGT	CC	TATTATT	GGA	GAAGT	ACTAG	AAGCA	ACTGG	TTATATT	TGT	5520
	CTAACAG	GA	TGGTCAT	TAC	ACCTATT	ACG	ATAAC	CGTAA	TGGATT	AGT	5580
	AAAAGG	CTTA	CCATTT	TAAAT	AGAGACT	TAT	TACAGG	TAC	AACAGC	TATT	5640
20	TCGTTGT	CA	GAAACC	ACAA	GCAGAG	GAAA	GACCTA	AGAA	TTGTG	CTTT	5700
	CAGCAT	CTCC	TAAGATT	GTA	CAAGA	ACAAA	AATTGT	TGGC	TATTG	AAAGT	5760
	ATGCTCT	TAC	TGAATTT	GGA	AGATAT	GCTG	ACATGT	TCTT	TATGG	CTGGA	5820
	TTAGGTT	GCT	GCTTGA	AGTC	TTTAA	ATAT	TGCTG	TTTT	ATTTAT	GTGT	5880
25	CTAAGAT	GCC	TAAAGT	TAAA	GTCAA	ACCAC	CTCTT	GCATT	TAAAG	TTTT	5940
	TCAGA	ACGCT	CAATTAC	ATG	AGACA	AATTGA	ACAA	ACCCTC	TGCTG	GCGT	6000
	TAGTTTT	TATT	GTTGAT	AGCA	ATATA	AATT	TCTTTT	TATT	GTTTGT	CAGT	6060
	TGCATA	AAAT	AACATG	TAA	GGTGT	GTAC	AGGCAT	AATA	AAATT	CTAGT	6120
30	CTGCAGT	CTG	TGGCA	ACTCT	ATTTT	ATGCA	AAGCCT	GTTT	GGCTT	CCTTAT	6180
	CTGATTT	TCA	ACATCT	CCAA	GTTACT	TGGG	ATTTCA	AATC	TGACCC	ACTA	6240
	TGGTACA	AAT	GTCTT	ACTTT	GCATT	CCTGG	CTGTTT	TGG	TAATA	ACTAT	6300
35	TTCTTAT	GTA	TTTTGT	ATCT	CAGTAC	CTCA	ACCTTT	GGCT	TTCTT	ATTTT	6360
	AGTACAG	TTG	GTTTTG	GCAT	GTTGT	CAACT	TTGAAT	CCAT	CTCAG	CTGAG	6420
	TAGTTAT	AGT	GGTTA	AGGCA	GTTCT	CGCCC	TTAA	ACATAT	TGTTTT	CGCA	6480
	CGTCTT	GCAA	AACGTG	CTCT	AGGACT	GCAA	GGCAG	ACACG	TATTC	CCTATT	6540
40	TTAATGG	TTC	AATGA	AGACT	GTTTAT	GTTC	ATGCTA	AATGG	TACTGG	TAAA	6600
	AACACA	AATTT	TTATTG	TAA	AACTGT	GATT	CTTATG	GCTT	TGAAA	ACACA	6660
	ACGAA	ATTGT	ACGTG	ATCTC	AGTA	AATAGT	TTAA	ACAAAC	TGTTT	ACGCC	6720
	CTCAT	CAAGA	AGTCAC	AAAA	GTTGA	ATGTT	CAGATG	GCTT	TTAC	AGATTT	6780
45	ATGAAT	TAC	TTCAT	ATGAC	TATGAT	GTAA	AACACA	AGAA	ATAC	AGTAGT	6840
	TCAAG	AGCAT	GCTCT	TGCTT	GATGAC	TTCA	TTGTG	TACAG	TCCAT	CTGGT	6900
	CAAA	TGTTAG	AAATG	CCTGT	GTTT	ACTTTT	CACA	ACTTAT	TGGT	AAGCCT	6960
	TTAAC	AGTGA	TTTG	CTTGAA	GACCT	CTCTG	TAGAT	TTTTAA	AGGGG	CACTT	7020
50	AAAAGA	ATGT	AATTA	AGAA	TCTTT	CAATG	TTGAT	GTCTC	AGAAT	GCAA	7080
	AATGT	TACAG	GGCTT	GCAAT	CTAAT	GTTT	CATTTT	CTAC	ATTTG	AAATG	7140
	ATGCT	CATAG	GTTT	GGTATT	CTGAT	TACTG	ATCGT	TCTTT	TAACA	AATTT	7200
	AAGTTA	AGCC	TGTTT	CATCT	GGTGT	GTCGG	CCATG	GACAT	TGGT	AAGTGT	7260
55	ATGCTA	AGAT	TGTTA	ATGCT	AAAGT	TTTTAA	CTCA	ACGTGG	TAAA	AGTGT	7320
	GCCAG	GATTT	TGCTG	CACTT	AGCTC	AACTG	CTCAG	AAAGT	TTTGG	TAAA	7380
	AAGA	AGGTGT	CAACTT	TTCAT	CTCAC	ATTTA	ATGCT	GTAGG	TTCAG	ATGAT	7440

60

ES 2 170 622 B1

5 ATGAAAGATT CACTGAATCT GTGTCTCCAA AAAGTGGTTC AGGCTTTTTC GATGTAATTA 7500  
 CACAGCTTAA ACAAATTGTG ATTTTGGTTT TTGTTTTTAT CTTTATTTGT GGTTTGTGCT 7560  
 CTGTTTACAG TGTTGCTACA CAGTCCCTACA TTGAATCTGC TGAAGGCTAT GACTACATGG 7620  
 TTATTAAGAA TGGAATTGTT CAACCTTTTG ACGATACCAT TTCATGTGTT CATAACACTT 7680  
 ATAAAGGATT CGGTGACTGG TTTAAAGCTA AGTATGGTTT TATCCCTACT TTTGGTAAAT 7740  
 10 CATGTCCAAT TGTTGTAGGA ACTGTTTTTG ATCTTGAAAA TATGAGACCA ATTCCTGACG 7800  
 TGCCTGCATA TGTTTCTATT GTGGGTAGAT CTCTTGTTTT CGCTATTAAT GCTGCTTTTG 7860  
 GTGTTACTAA TATGTGCTAT GATCATACTG GCAATGCAGT TAGTAAGGAC TCTTACTTTG 7920  
 ATACTTGTGT GTTTAATACT GCGTGCACCA CTCCTACAGG TCTTGGTGGT ACAAATTGTAT 7980  
 15 ATTGTGCAA GCAAGGTTTA GTTGAAGGTG CTAAGCTCTA TAGTGATCTT ATGCCAGACT 8040  
 ATTATTATGA GCATGCTAGT GGTAACATGG TTAAATTGCC AGCAATTATT AGAGGACTTG 8100  
 GTCTACGTTT TGTGAAAACA CAGGCTACAA CTTATTGTAG AGTGGGAGAG TGCATTGATA 8160  
 GTAAAGCTGG TTTTGTCTTT GGTGGCGATA ACTGGTTTGT CTACGACAAT GAGTTTGGCA 8220  
 20 ATGGATACAT CTGTGGTAAAT TCTGTGCTAG GATTCCTTAA GAATGTCTTC AAACCTCTTA 8280  
 ACTCTAACAT GTCTGTGGTA GCTACATCTG GTGCGATGCT TGTTAACATT ATTATTGCAT 8340  
 GCTTAGCTAT TGCAATGTGT TATGGTGTTC TTAAGTTTAA GAAGATTTTT GGTGATTGTA 8400  
 CTTTCCTCAT TGTTATGATC ATTGTCACCC TTGTTGTGAA CAATGTGTCT TATTTTGTCA 8460  
 25 CTCAAAAAC GTTCTTTATG ATCATCTACG CCATTGTTTA CTATTTTATA ACAAGAAAAC 8520  
 TTGCATACCC AGGCATTCTT GATGCTGGGT TTATTATTGC TTATATTAAT ATGGCTCCAT 8580  
 GGTACGTGAT TACCGCATAT ATCCTAGTTT TCCTCTATGA CTCACTCCCT TCACTGTTTA 8640  
 AACTTAAAGT TTCAACAAAT CTTTTTGAAG GTGATAAATT TGTGGGTAAC TTTGAATCTG 8700  
 30 CTGCTATGGG TACTTTTGTT ATTGACATGC GTTCATATGA AACTATTGTT AATTCTACTT 8760  
 CTATTGCTAG AATTAATCA TATGCTAACA GCTTCAATAA ATATAAGTAC TACACAGGTT 8820  
 CAATGGGAGA AGCTGACTAC AGAATGGCTT GCTATGCTCA TCTTGGTAAA GCTCTTATGG 8880  
 35 ACTATTCTGT TAATAGAACA GACATGCTTT ACACACCTCC TACTGTTAGT GTTAATTCTA 8940  
 CACTTCAGTC AGGTTTGC GG AAAATGGCAC AGCCTAGTGG TCTTGTAGAG CCTTGCATTG 9000  
 TAAGAGTTT CTATGGTAAC AATGTGCTTA ATGGTTTATG GTTAGGAGAT GAAGTCATTT 9060  
 GCCCTAGACA TGTTATTGCT AGTGATACCA CACGTGTTAT CAACTATGAA AATGAAATGT 9120  
 40 CTAGTGTGAG ACTTCACAAC TTTTCAGTTT CTAAGAATAA TGTGTTTTTG GGTGTGTGT 9180  
 CTGCCAGATA TAAGGGTGTG AATCTTGTAC TTAAAGTCAA CCAGGTTAAT CCTAACACAC 9240  
 CAGAACATAA ATTTAAGTCT ATTTAAAGCTG GTGAAAGTTT TAACATTCTT GCTTGTATG 9300  
 AAGCATGTCC TGGCAGTGT TATGGTGTCA ACATGAGAAG TCAAGGTACC ATTTAAAGGAT 9360  
 45 CTTTTATAGC TGTTACTTGT GGATCAGTAG GTTATGTGTT AGAAAATGGA ATTCTCTATT 9420  
 TTGTATACAT GCATCACTTA GAAGTGGAA ATGGCTCGCA TGTTGGTTCC AATTTTGAAG 9480  
 GAGAAATGTA CGGTGGTTAT GAAGATCAAC CTAGCATGCA ATTGGAAGGT ACTAATGTCA 9540  
 TGTCATCAGA TAATGTGGTT GCATTCTAT ATGCTGCACT TATCAATGGT GAAAGGTGGT 9600  
 50 TTGTTACAAA CACATCGATG TCATTAGAAT CATACAATAC ATGGGCCAAA ACTAACAGTT 9660  
 TCACAGAACT TTCTTCAACT GATGCTTTTA GCATGTTGGC TGCAAAAACCT GGTCAAAGTG 9720  
 TTGAGAAAT ACTAGATAGC ATCGTAAGAC TCAACAAGGG TTTTGGAGGT CGTACTATAC 9780  
 TTTCTTATGG CTCATTGTGT GACGAGTTCA CTCCAAGTGA AGTCATAAGG CAAATGTATG 9840  
 55 CTGTAATCT TCAGGCTGGT AAAGTAAAT CTTTCTTCTA CCCTATTATG ACTGCAATGA 9900  
 CAATTCTCTT TGCTTTTGG CTTGAATCT TTATGTACAC ACCCTTCACT TGGATTAATC 9960  
 CAACTTTTGT TAGCAATTGTA FTGGGTGTA CAACTTTGAT CTCGACGGT TTTGTCTCTG 10020

60

# ES 2 170 622 B1

	GCATCAAACA	TAAGATGTTG	TTCTTTATGT	CTTTTGTCTT	TCCTAGTGTT	ATCCTTGTGA	10080
5	CAGCACACAA	TTGTTTCTGG	GACTTTTCTT	ACTATGAAAG	TCTTCAGTCA	ATTGTTGAGA	10140
	ATACTAACAC	TATGTTTTTG	CCTGTTGACA	TGCAAGGTGT	CATGCTCACA	GTGTTTTGCT	10200
	TTATTGTCTT	TGTTACATAT	AGTGTTAGAT	TCTTCACTTG	CAAACAATCA	TGGTTCTCAC	10260
	TTGCTGTGAC	AACTATTCTT	GTGATCTTTA	ACATGGTTAA	AATCTTTGGA	ACATCTGATG	10320
10	AACCATGGAC	TGAAAACCAA	ATTGCTTTCT	GCTTTGTGAA	CATGCTTACT	ATGATTGTCA	10380
	GTCTTACTAC	AAAGGATTGG	ATGTTTGTCA	TTGCATCATA	CAGAATTGCA	TATTATATTG	10440
	TTGTATGTGT	AATGCCATCT	GCTTTTGTAT	CTGACTTTGG	GTTTATGAAG	TGTATTAGCA	10500
	TTGTTTACAT	GGCGTGCGGT	TATTTGTTTT	GTTGCTATTA	TGGCATTCTT	TATTGGGTTA	10560
15	ACAGATTTAC	ATGCATGACT	TGTGGTGTTC	ATCAATTCAC	TGTGTCTGCA	GCTGAACTTA	10620
	AATACATGAC	CGCTAACAAAC	CTTCTGACAC	CTAAGAACGC	ATATGACGCT	ATGATTCTTA	10680
	GTGCTAAATT	GATTGGTGTG	GGAGGTAAGA	GAAACATCAA	AATTTCAACT	GTACAGTCAA	10740
	AACTTACAGA	GATGAAATGT	ACCAATGTTG	TCTTGCTTGG	TCTTTTATCT	AAAATGCATG	10800
20	TCGAGTCTAA	CTCAAAAGAG	TGGAAGTATT	GTGTTGGACT	ACACAATGAG	ATAAACCTTT	10860
	GTGACGATCC	TGAAATCGTT	CTTGAGAAAC	TGTTAGCTCT	TATTGCATTC	TTCTTGTCCA	10920
	AACATAACAC	TTGTGACCTT	AGCGAACTTA	TTGAATCATA	CTTTGAGAAC	ACCACCATAC	10980
	TCCAGAGTGT	GGCTTCAGCT	TATGCTGCAT	TGCCTAGCTG	GATTGCACTT	GAAAAAGCTC	11040
25	GCGCTGATCT	TGAAGAGGCT	AAGAAAAATG	ATGTTAGCCC	TCAAATTTTG	AAGCAGCTTA	11100
	CTAAAGCATT	TAACATTGCC	AAGAGTGATT	TTGAGCGCGA	AGCATCAGTG	CAAAGAAAC	11160
	TCGACAAAAT	GGCTGAGCAG	GCTGCAGCTA	GTATGTATAA	AGAAGCACGA	GCTGTGGACA	11220
	GAAAGTCAA	GATTGTTTCT	GCTATGCATA	GCCTACTTTT	TGGTATGCTT	AAGAACTTG	11280
30	ATATGTCCAG	TGTCAACACT	ATTATTGACC	AGGCTCGTAA	TGGTGTCTA	CCTTTAAGTA	11340
	TCATTTCCAGC	TGCATCAGCT	ACAAGACTTG	TTGTTATTAC	ACCTAGCCTT	GAAGTGTTTT	11400
	CCAAGATTAG	GCAAGAAAAC	AATGTTTCATT	ATGCTGGTGC	TATTTGGACT	ATTGTTGAAG	11460
	TTAAAGATGC	TAATGGTTCA	CATGTACATC	TTAAGGAAGT	CACCGCTGCT	AATGAATTAA	11520
35	ACCTTACTTG	GCCATTGAGC	ATTACTTGTG	AGAGAACCAC	AAAGCTTCAG	AACAATGAAA	11580
	TTATGCCAGG	TAAACTTAAA	GAAAGAGCTG	TCAGAGCGTC	AGCAACTCTT	GATGGTGAAG	11640
	CTTTCGGCAG	TGGAAAGGCT	CTTATGGCAT	CTGAAAGTGG	AAAAGCTTT	ATGTATGCAT	11700
	TTATAGCCTC	AGACAACAAT	CTTAAGTATG	TTAAGTGGGA	GAGCAATAAT	GATATTATAC	11760
40	CTATTGAACT	TGAAGCTCCA	TTGCGTTTCT	ATGTTGACGG	CGCTAATGGT	CCTGAAGTCA	11820
	AGTATTTGTA	TTTTGTCAAG	AATTTAAACA	CTCTTAGACG	TGGTGCCGTT	CTTGGTTATA	11880
	TCGGTGCAAC	AGTTCGTCTG	CAAGCTGGTA	AACCCACTGA	ACATCCATCT	AACAGTAGTT	11940
45	TATTGACATT	GTGTGCTTTT	TCACCTGATC	CTGCTAAAGC	ATATGTTGAT	GCTGTTAAGA	12000
	GAGGCATGCA	ACCAGTTAAT	AACTGTGTAA	AAATGCTCTC	AAATGGTGTG	GGTAATGGTA	12060
	TGGCTGTTAC	AAACGGTGTG	GAAGCTAACA	CACAACAGGA	CTCTTATGGT	GGTGCTTCAG	12120
	TTTGTATTTA	TTGCAGATGC	CATGTTGAAC	ATCCTGCTAT	TGATGGATTA	TGCCGCTACA	12180
50	AAGGTAAGTT	CGTGCAAATA	CCAAGTGGCA	CACAAGATCC	AATTCGGTTC	TGTATTGAAA	12240
	ATGAAGTTTG	TGTTGTCTGT	GGTTGTTGGC	TTAACAATGG	TTGCATGTGC	GATCGTACTT	12300
	CTATGCAGAG	TTTTACTGTT	GATCAAAGTT	ATTTAAACGA	GTGCGGGGTT	CTAGTGCAGC	12360
	TCGACTAGAA	CCCTGCAATG	GTACTGATCC	AGACCATGTT	AGTAGAGCTT	TTGACATCTA	12420
55	CAACAAAGAT	GTTGCGTGTA	TTGGTAAATT	CCTTAAGACG	AATTGTTCAA	GATTTAGGAA	12480
	TTTGGACAAA	CATGATGCCT	ACTACATTGT	CAAACGTTGT	ACAAAGACCG	TTATGGACCA	12540
	TGAGCAAGTC	TGTTATTAACG	ATCTTAAAGA	TTCTGGTST	GTGCTGAGC	ATGACTTCTT	12600

60

ES 2 170 622 B1

	CACATATAAA	GAGGGTAGAT	GTGAGTTCGG	TAATGTTGCA	CGTAGGAATC	TTACAAAGTA	12660
5	CACAATGATG	GATCTTTGTT	ACGCTATCAG	AAATTTTGAT	GAAAAGAACT	GTGAAGTTCT	12720
	CAAAGAAATA	CTCGTGACAG	TAGGTGCTTG	CACTGAAGAA	TTCTTTGAAA	ATAAAGATTG	12780
	GTTTGATCCA	GTTGAAAATG	AAGCCATACA	TGAAGTTTAT	GCAAAACTTG	GACCCATTGT	12840
	AGCCAATGCT	ATGCTTAAAT	GTGTTGCTTT	TTGCGATGCG	ATAGTGGAAA	AAGGCTATAT	12900
10	AGGTGTTATA	ACACTTGACA	ACCAAGATCT	TAATGGCAAT	TTCTACGATT	TCGGCGATTT	12960
	CGTGAAGACT	GCTCCGGGTT	TTGGTTGCGC	TTGTGTTACA	TCATATTATT	CTTATATGAT	13020
	GCCTTTAATG	GGGATGACTT	CATGCTTAGA	GTCTGAAAAC	TTTGTGAAAA	GTGACATCTA	13080
	TGGTCTGAT	TATAAGCAGT	ATGATTTACT	AGCTTATGAT	TTTACCGAAC	ATAAGGAGTA	13140
15	CCTTTTCCAA	AAATACTTTA	AGTACTGGGA	TCGCACATAT	CACCCAAATT	GTTCTGATTG	13200
	TACTAGTGAC	GAGTGTATTA	TTCATTGTGC	TAATTTTAAAC	ACATTGTTTTT	CTATGACAAT	13260
	ACCAATGACA	GCTTTTGGAC	CACTTGTCCG	TAAAGTTCAT	ATTGATGGTG	TACCAGTAGT	13320
20	TGTTACTGCA	GTTTACCATT	TCAAACAAC	TGGTATAGTA	TGGAATCTTG	ATGTAAAATT	13380
	AGACACAATG	AAGTTGAGCA	TGACTGATCT	TCTTAGATTT	GTCACAGATC	CAACACTTCT	13440
	TGTAGCATCA	AGCCCTGCAC	TTTTAGACCA	GCGTACTGTC	TGTTTCTCCA	TTGCAGCTTT	13500
	GAGTACTGGT	ATTACATATC	AGACAGTAAA	ACCAGGTCAC	TTAACAAAG	ATTTCTACGA	13560
25	TTTCATAACA	GAGCGTGGAT	TCTTTGAAGA	GGGATCTGAG	TTAACATTAA	AACATTTTTT	13620
	CTTTGCACAG	GGTGGTGAAG	CTGCATGAC	AGACTTCAAT	TATTATCGCT	ACAATAGAGT	13680
	CACAGTACTT	GATATTTGCC	AAGCTCAATT	TGTTTACAAA	ATAGTTGGCA	AGTATTTTGA	13740
	ATGTTATGAC	GGTGGGTGCA	TTAATGCTCG	TGAAGTTGTT	GTTACAAACT	ATGACAAGAG	13800
30	TGCTGGCTAT	CCTTTGAACA	AATTTGGTAA	AGCTAGACTT	TACTACGAAA	CTCTTTCATA	13860
	TGAAGAGCAG	GATGCACTTT	TTGCTTTAAC	AAAGAGAAAT	GTTTTACCCA	CAATGACTCA	13920
	AATGAATTTG	AAATACGCTA	TTTCTGGTAA	GGCAAGAGCT	CGTACAGTAG	GAGGAGTTTC	13980
35	ACTTCTTTCT	ACCATGACTA	CGAGACAATA	TCATCAGAAG	CATTTGAAGT	CAATTGCTGC	14040
	AACACGCAAT	GCTACTGTGG	TCATTGGTTC	AACCAAGTTT	TATGGTGGTT	GGGACAATAT	14100
	GCTTAAAAAT	TTAATGCGTG	ATGTTGATAA	TGTTTGTGTT	ATGGGATGGG	ACTATCCTAA	14160
	GTGTGACCGT	GCTTTACCTA	ATATGATTAG	AATGGCTTCT	GCCATGATAT	TAGGTTCTAA	14220
40	GCATGTTGGT	TGTTGTACAC	ATAATGATAG	GTTCTACCGC	CTCTCCAATG	AGTTAGCTCA	14280
	AGTACTCACA	GAAGTTGTGC	ATTGCACAGG	TGTTTTTTAT	TTTAAACCTG	GTGGTACAAC	14340
	TAGCGGTGAT	GGTACTACAG	CATATGCTAA	CTCTGCTTTT	AACATCTTTC	AAGCTGTTTC	14400
	TGCTAATGTT	AATAAGCTTT	TGGGGGTTGA	TTCAAACGCT	TGTAACAACG	TTACAGTAAA	14460
45	ATCCATACAA	CGTAAAATTT	ACGATAATTG	TTATCGTAGT	AGCAGCATTG	ATGAAGAATT	14520
	TGTTGTTGAG	TACTTTAGTT	ATTTGAGAAA	ACACTTTTCT	ATGATGATTT	TATCTGATGA	14580
	TGGAGTTGTG	TGCTACAACA	AAGATTATGC	GGATTTAGGT	TATGTAGCTG	ACATTAATGC	14640
50	TTTTAAAGCA	ACACTTTTATT	ACCAGAATAA	CGTCTTTATG	TCCACTTCTA	AGTGTGGGGT	14700
	AGAACCAGAT	CTTAGTGTTG	GACCACATGA	ATTTTGTTC	CAGCATACAT	TGCAGATTGT	14760
	TGGGCCTGAT	GGAGACTACT	ATCTTCCCTA	TCCAGACCCG	TCCAGAATTT	TATCAGCTGG	14820
	TGTGTTTGTT	GATGACATAG	TTAAAACAGA	CAATGTTATT	ATGTTAGAAC	GTTACGTGTC	14880
55	ATTGGCTATT	GACGCATACC	CGCTCACAAA	ACACCCTAAG	CCTGCTTATC	AAAAAGTGTT	14940
	TTACTACTCTA	CTAGATTGGG	TTAAACATCT	ACAGAAAAAT	TTGAATGCAG	GTGTTCTTGA	15000
	TTGCTTTTCA	GTGACAATGT	TAGAGGAAGG	TCAAGATAAG	TTCTGGAGTG	AAGAGTTTTA	15060
	CGCTAGCCCTC	TATGAAAAGT	CCACTGTCTT	GCAAGCTGCA	GGCATGTGTG	TAGTATGTGG	15120
60	TTGCGAAACT	GTACTTGGTT	STGGAGACTG	TCTTAGGAGA	CCACTTTTAT	GCACGAAATG	15180

# ES 2 170 622 B1

	TGCTTACGAC	CATGTTATGG	GAACAAAGCA	TAAATTCATT	ATGTCTATCA	CACCATATGT	15240
5	GTGTAGTTTT	AATGGTTGTA	ATGTCAATGA	TGTTACAAAG	TTGTTTTTAG	GTGGTCTTAG	15300
	TTATTATTGT	ATGAACCACA	AACCACAGTT	GTCATTCCCA	CTCTGTGCTA	ATGGCAACGT	15360
	TTTTGGTCTA	TATAAAAGTA	GTGCAGTCGG	CTCAGAGGCT	GTTGAAGATT	TCAACAAACT	15420
	TGCAGTTTCT	GACTGGACTA	ATGTAGAAGA	CTACAAACTT	GCTAACAATG	TCAAGGAATC	15480
10	TCTGAAAATT	TTCGCTGCTG	AAACTGTGAA	AGCTAAGGAG	GAGTCTGTTA	AATCTGAATA	15540
	TGCTTATGCT	GTATTAAGG	AGGTTATCGG	CCCTAAGGAA	ATTGTAAGCT	AATGGGAAGC	15600
	TTCTAAGACT	AAGCCTCCAC	TTAACAGAAA	TTCAGTTTTC	ACGTGTTTTT	AGATAAGTAA	15660
	GGATACTAAA	ATTCAATTAG	GTGAATTTGT	GTTTGAGCAA	TCTGAGTACG	GTAGTGATTC	15720
15	TGTTTATTAC	AAGAGCACGA	GTAAGTACAA	ATTGACACCA	GGTATGATTT	TTGTGTTGAC	15780
	TTCTCATAAT	GTGAGTCCTC	TTAAAGCTCC	AATTTTAGTC	AACCAAGAAA	AGTACAATAC	15840
	CATATCTAAG	CTCTATCCTG	TCTTTAATAT	AGCGGAGGCC	TATAATACAC	TGGTTCCTTA	15900
20	CTACCAAATG	ATAGGTAAGC	AAAAATTTAC	AACTATCCAA	GGTCCCTCCTG	GTAGCGGTAA	15960
	ATCTCATTGT	GTTATAGGTT	TGGGTTTGTA	TTACCCTCAG	GCGAGAATAG	TCTACACTGC	16020
	ATGTTCTCAT	GCGGCTGTAG	ACGCTTTATG	TGAAAAAGCA	GCCAAAAACT	TCAATGTTGA	16080
	TAGATGTTCA	AGGATAATAC	CTCAAAGAAT	CAGAGTTGAT	TGTTACACAG	GCTTTAAGCC	16140
25	TAATAACACC	AATGCGCAGT	ACTTGTTTTG	TACTGTTAAT	GCTCTACCAG	AAGCAAGTTG	16200
	TGACATTGTT	GTAGTTGATG	AGGTCTCTAT	GTGACTAAT	TATGATCTTA	GTGTCATAAA	16260
	TAGCCGACTG	AGTTACAAAC	ATATTGTTTA	TGTTGGAGAC	CCACAGCAGC	TACCAGCTCC	16320
	TAGAACTTTG	ATTAATAAGG	GTGACTTCA	ACCGCAGGAT	TACAATGTTG	TAACCAAAG	16380
30	AATGTGCACA	CTAGGACCTG	ATGCTTTTTT	GCATAAATGT	TACAGGTGCC	CAGCTGAAAT	16440
	TGTTAAGACA	GTCTCTGCAC	TTGTTTATGA	AAATAAATTT	GTACCTGTCA	ACCCAGAATC	16500
	AAAGCAGTGC	TTCAAAATGT	TTGTAAAAGG	TCAGGTTGAG	ATTGAGTCTA	ACTCTTCTAT	16560
35	AAACAACAAG	CAACTAGAGG	TTGTCAAGGC	CTTTTTAGCA	CATAATCCAA	AATGGCGTAA	16620
	AGCTGTTTTT	ATCTCACCTT	ATAATAGTCA	AAATTATGTT	GCTCGGCGTC	TTCTTGGTTT	16680
	GCAAACGCAA	ACTGTGGATT	CCGCTCAGGG	TAGTGAGTAT	GATTACGTCA	TCTACACACA	16740
	GACCTCCGAT	ACACAGCATG	CTACTAATGT	TAACAGATTT	AATGTTGCCA	TTACGAGAGC	16800
40	AAAGGTTGGT	ATACTTTGTA	TCATGTGTGA	TAGAACTATG	TATGAGAATC	TTGATTTCTA	16860
	TGAACTCAA	GATTCAAAGA	TTGGTTTACA	AGCAAACCT	GAAACTTGTTG	GTTTATTTAA	16920
	AGATTGTTTCG	AAGAGCGAAC	AATACATACC	ACCTGCTTAT	GCAACGACAT	ATATGAGCTT	16980
	ATCTGATAAT	TTTAAGACAA	GTGATGGTTT	AGCTGTTAAC	ATCGGTACAA	AAGATGTTAA	17040
45	ATATGCTAAT	GTCATCTCAT	ATATGGGATT	CAGGTTTGAA	GCCAACATAC	CAGGCTATCA	17100
	CACACTATTC	TGCACGCGAG	ATTTTGCTAT	GCGTAATGTT	AGAGCATGGC	TTGGGTTTGA	17160
	CGTTGAAGGT	GCACATGTCT	GTGGTGATAA	TGTTGGAAT	AATGTACCAT	TACAGCTGGG	17220
50	TTTCTCAAAC	GGTGTGGATT	TTGTAGTGCA	AACTGAAGGA	TGTGTTATTA	CTGAAAAAGG	17280
	TAATAGCATT	GAGGTTGTAA	AAGCACGAGC	ACCACCAGGT	GAGCAATTTG	CACACTTGAT	17340
	TCCGCTTATG	AGAAAGGGTC	AACCTTGCCA	CATTGTTAGA	CGCCGTATAG	TGCAGATGGT	17400
	CTGTGACTAT	TTTGATGGCT	TATCAGACAT	TCTGATCTTT	GTGCTTTGGG	CTGGTGGTCT	17460
55	TGAACTTACA	ACTATGAGAT	ACTTTGTTAA	AATTGGAAGA	CCACAAAAAT	GTGAATGCGG	17520
	CAAAAGTGCA	ACTTGTATA	GTAAGTCTCA	ATCTGTTTAT	GCTTGCTTCA	AGCATGCATT	17580
	AGGATGTGAT	TATTTATATA	ACCCTTACTG	CATTGACATA	CAGCAATGGG	GTTACACAGG	17640
	ATCTTTGAGC	ATGAATCATC	ATGAAGTTTG	CAACATTCAT	AGAAATGAGC	ATGTAGCTAG	17700
60	TGGTGATGCT	ATCATGACTA	GATGCTCGC	TATACATGAC	TGTTTTGTCA	AACCGTGTGA	17760



# ES 2 170 622 B1

	TTGGTCAATT	GTGTACCCTT	TTATTGACAA	TGAAGAAAAG	ATCAATAAAG	CTGGTCGCAT	17820
5	AGTGCAGTCA	CATGTCATGA	AAGCTGCTCT	GAAGATTTTT	AATCCTGCTG	CAATTCACGA	17880
	TGTGGGTAAT	CCAAAAGGCA	TCCGTTGTGC	TACAACACCA	ATACCATGGT	TTTGTATGA	17940
	TCGTGATCCT	ATTAATAACA	ATGTTAGATG	TCTGGATTAT	GACTATATGG	TACATGGTCA	18000
	AATGAATGGT	CTTATGTTAT	TTTGGAACTG	TAATGTAGAC	ATGTACCCAG	AGTTTTCAAT	18060
10	TGTTTGTAGA	TTTGATACTC	GCACTCGCTC	TAAATTGTCT	TTAGAAGGTT	GTAATGGTGG	18120
	TGCATTGTAT	GTTAATAACC	ATGCTTTCCA	CACACCAGCT	TATGATAGAA	GAGCTTTTGC	18180
	TAAGCTTAAA	CCTATGCCAT	TCTTTTACTA	TGATGATAGT	AATTGTGAAC	TTGTTGATGG	18240
	GCAACCTAAT	TATGTACCAC	TTAAGTCAAA	TGTTTGCATA	ACAAAATGCA	ACATTGGTGG	18300
15	TGCTGTCTGC	AAGAAGCATG	CTGCTCTTTA	CAGAGCGTAT	GTTGAGGATT	ACAACATTTT	18360
	TATGCAGGCT	GGTTTTACAA	TATGGTGTCC	TCAAAACTTT	GACACCTATA	TGCTTTGGCA	18420
	TGGTTTTGTT	AATAGCAAAG	CACCTCAGAG	TCTAGAAAAT	GTGGCTTTTA	ATATCGTTAA	18480
	GAAAGGTGCC	TTCACCGGTT	TAAAAGGTGA	CTTACCAACT	GCTGTTATTG	CTGACAAAAT	18540
20	AATGGTAAGA	GATGGACCTA	CTGACAAATG	TATTTTTACA	AATAAGACTA	GTTTACCTAC	18600
	AAATGTAGCT	TTTGAGTTAT	ATGCAAAAACG	CAAACCTGGA	CTCACACCTC	CATTAACAAT	18660
	ACTTAGGAAT	TTAGGTGTTG	TCGCAACATA	TAAGTTTGTG	TTGTGGGATT	ATGAAGCTGA	18720
	ACGTCCTTTC	TCAAATTTCA	CTAAGCAAGT	GTGTTCTTAC	ACTGATCTTG	ATAGTGAAGT	18780
25	TGTAACATGT	TTTGATAATA	GTATTGCTGG	TTCTTTTGAG	CGTTTTACTA	CTACAAGAGA	18840
	TGCAGTGCCT	ATTTCTAATA	ACGCTGTGAA	AGGGCTTAGT	GCCATTAAT	TACAATATGG	18900
	CCTTTTGAAT	GATCTACCTG	TAAGTACTGT	TGGAATAAAA	CCTGTACACAT	GGTATATCTA	18960
	TGTGCGCAAG	AATGGTGAGT	ACGTCGAACA	AATCGATAGT	TACTATACAC	AGGGACGTAC	19020
30	TTTTGAAACC	TTCAAACCTC	GTAGTACAAT	GGAAGAAGAT	TTTCTTAGTA	TGGATACTAC	19080
	ACTCTTCATC	CAAAAGTATG	GTCTTGAGGA	TTATGGTTTT	GAACACGTTG	TATTTGGAGA	19140
	TGTCTCTAAA	ACTACCATTG	GTGGTATGCA	TCTTCTTATA	TCGCAAGTGC	GCCTTGCAAA	19200
	AATGGGTTTG	TTTTCCGTTT	AAGAATTTAT	GAATAATTTCT	GACAGTACAC	TGAAAAGTTG	19260
35	TTGTATTACA	TATGCTGATG	ATCCATCTTC	TAAGAATGTG	TGCACTTATA	TGGACATACT	19320
	CTTGGACGAT	TTTGTGACTA	TCATTAAGAG	CTTAGATCTT	AATGTTGTGT	CCAAAGTTGT	19380
	GGATGTCATT	GTAGATTGTA	AGGCATGGAG	ATGGATGTTG	TGGTGTGAGA	ATTCACATAT	19440
	TAAAACCTTC	TATCCACAAC	TCCAATCTGC	TGAATGGAAT	CCC GGCTATA	GCATGCCTAC	19500
40	ACTGTACAAA	ATCCAGCGTA	TGTGTCTCGA	ACGGTGTAAT	CTCTACAATT	ATGGTGCACA	19560
	AGTGAAATTA	CCTGTAGGCA	TTACTACTAA	GTTTCGTTAAG	TATACTCAGT	TGTGTCAATA	19620
	CCTTAACACT	ACTACATTGT	GTGTACCACA	CAAAATGCGT	GTATTGCATT	TAGGAGCTGC	19680
	TGGTGCATCT	GGTGTGCTC	CTGGTAGTAC	TGTATTAAGA	AGATGGTTAC	CAGATGATGC	19740
45	CATATTGGTT	GATAATGATT	TGAGAGATTA	CGTTTCCGAC	GCAGACTTCA	GTGTTACAGG	19800
	TGATTGTACT	AGTCTTTACA	TCGAAGACAA	GTTTGATTTG	CTCGTCTCTG	ATTTATATGA	19860
	TGGCTCCACA	AAATCAATTG	ACGGTGAAAA	CACGTCGAAA	GATGGTTTCT	TTACTTATAT	19920
	TAATGGTTTT	ATTAAAGAGA	AACTGTCACT	TGGTGGATCT	GTTGCCATTA	AAATCACGGA	19980
50	ATTTAGTTGG	AATAAAGATT	TATATGAATT	GATTCAAAGA	TTTGAGTATT	GGACTGTGTT	20040
	TTGTACAAGT	GTTAACACGT	CATCATCAGA	AGGCTTTCTG	ATTGGTATTA	ACTACTTAGG	20100
	ACCATACTGT	GACAAAGCAA	TAGTAGATGG	AAATATAATG	CATGCCAATT	ATATATTTTG	20160
55	GAGAACTCT	ACAATTATGG	CTCTATCACA	TAACCTCAGTC	CTAGACACTC	CTAAATTCAA	20220
	GTGTCGTTGT	AACAACGCAC	TTATTGTTAA	TTTAAAAGAA	AAAGAATTGA	ATGAAATGGT	20280
	CATTGGATTA	CTAASGAAGG	CTAAGTTGCT	CATLAGAAAT	AATGGTAAGT	FACTAAACTT	20340

60

# ES 2 170 622 B1

	TGGTAACCAC	TCGTAAACA	CACCATGAAA	AACTATTTG	TGGTTTTGGT	CGTAATGCCA	20400
5	TTGATTTATG	GAGACAATTT	TCCTTGTTCT	AAATTGACTA	ATAGAACTAT	AGGCAACCAG	20460
	TGGAATCTCA	TTGAAACCTT	CCTTCTAAAC	TATAGTAGTA	GGTACCACC	TAATTCAGAT	20520
	GTGGTGTTAG	GTGATTATTT	TCCTACTGTA	CAACCTTGGT	TTAATTGCAT	TCGCAATGAT	20580
	AGTAATGACC	TTTATGTTAC	ACTGGAAAAT	CTTAAAGCAT	TGTATTGGGA	TTATGCTACA	20640
10	GAAAATATCA	CTTGAATCA	CAGACAACGG	TTAAACGTAG	TCGTTAATGG	ATACCATAAC	20700
	TCCATCACAG	TTACAACAAC	CCGCAATTTT	AATTCTGCTG	AAGGTGCTAT	TATATGCATT	20760
	TGTAAGGGCT	CACCACCTAC	TACCACCACA	GAATCTAGTT	TGACTTGCAA	TTGGGGTAGT	20820
	GAGTGCAGGT	TAAACCATAA	GTTCCTATA	TGTCCTTCTA	ATTCAGAGGC	AAATTGTGGT	20880
15	AATATGCTGT	ATGGCCTACA	ATGGTTTGA	GATGAGGTTG	TTGCTATTTT	ACATGGTGCT	20940
	AGTTACCGTA	TTAGTTTTGA	AAATCAATGG	TCTGGCACTG	TCACATTTGG	TGATATGCGT	21000
	GCGACAACAT	TAGAAGTCGC	TGGCACGCTT	GTAGACCTTT	GGTGGTTTAA	TCCTGTTTAT	21060
20	GATGTCAGTT	ATTATAGGGT	TAATAATAAA	AATGGTACTA	CCGTAGTTTC	CAATTGCACT	21120
	GATCAATGTG	CTAGTTATGT	GGCTAATGTT	TTTACTACAC	AGCCAGGAGG	TTTATAACCA	21180
	TCAGATTTTA	GTTTTAATAA	TTGGTTCCTT	CTAACTAATA	GCTCCACGTT	GGTTAGTGGT	21240
	AAATTAGTTA	CCAAACAGCC	GTTATTAGTT	AATTGCTTAT	GGCCAGTCCC	TAGCTTTGAA	21300
25	GAAGCAGCTT	CTACATTTTG	TTTTGAGGGT	GCTGGCTTTG	ATCAATGTAA	TGGTGCTGTT	21360
	TTAAATAATA	CTGTAGACGT	CATTAGGTTT	AACCTTAATT	TTACTACAAA	TGTACAATCA	21420
	GGTAAGGGTG	CCACAGTGTT	TTCATTGAAC	ACAACGGGTG	GTGTCACTCT	TGAAATTTCA	21480
	TGTTATACAG	TGAGTGACTC	GAGCTTTTTT	AGTTACGGTG	AAATCCGTTT	CGGCCTAACT	21540
30	GATGGACCAC	GGTACTGTTA	CGTACACTAT	AATGGCACAG	CTCTTAAGTA	TTTAGGAACA	21600
	TTACCACCTA	GTGTCAAGGA	GATTGCTATT	AGTAAGTGGG	GCCATTTTTA	TATTAATGGT	21660
	TACAATTTCT	TTAGCACATT	TCCTATTGAT	TGTATATCTT	TTAATTTGAC	CACTGGTGAT	21720
35	AGTGACGTTT	TCTGGACAAT	AGCTTACACA	TCGTACTACTG	AAGCATTAGT	ACAAGTTGAA	21780
	AACACAGCTA	TTACAAAGGT	GACGTATTGT	AATAGTCACG	TTAATAACAT	TAAATGCTCT	21840
	CAAATTACTG	CTAATTTGAA	TAATGGATTT	TATCCTGTTT	CTTCAAGTGA	AGTTGGTCTT	21900
	GTCAATAAGA	GTGTTGTGTT	ACTACCTAGC	TTTTACACAC	ATACCATTGT	TAACATAACT	21960
40	ATTGGTCTTG	GTATGAAGCG	TAGTGGTTAT	GGTCAACCCA	TAGCCTCAAC	ATTAAGTAAC	22020
	ATCACACTAC	CAATGCAGGA	TCACAACACC	GATGTGTACT	GTATTCGTTC	TGACCAATTT	22080
	TCAGTTTATG	TTCATTCTAC	TTGCAAAAGT	GCTTTATGGG	ACAATATTTT	TAAGCGAAAC	22140
	TGCACGGACG	TTTTAGATGC	CACAGCTGTT	ATAAAAAC TG	GTACTTGTC	TTTCTCATT	22200
45	GATAAATTGA	ACAATTACTT	AACTTTTAAAC	AAGTTCTGTT	TGTCGTTGAG	TCCTGTTGGT	22260
	GCTAATTGTA	AGTTTGATGT	AGCTGCCCCG	ACAAGAACCA	ATGAGCAGGT	TGTTAGAAGT	22320
	TTGTATGTAA	TATATGAAGA	AGGAGACAAC	ATAGTGGGTG	TACCGTCTGA	TAATAGTGGT	22380
50	GTGCACGATT	TGTCAGTGCT	ACACCTAGAT	TCCTGCACAG	ATTACAATAT	ATATGGTAGA	22440
	ACTGGTGTTG	GTATTATTAG	ACAAACTAAC	AGGACGCTAC	TTAGTGGCTT	ATATTACACA	22500
	TCACTATCAG	GTGATTTGTT	AGGTTTTTAAA	AATGTTAGTG	ATGGTGTTCAT	CTACTCTGTA	22560
	ACGCCATGTG	ATGTAAGCGC	ACAAGCAGCT	GTTATTGATG	GTACCATAGT	TGGGGCTATC	22620
55	ACTTCCATTA	ACAGTGAAC	GTTAGGTCTA	ACACATTGGA	CAACAACACC	TAATTTTTAT	22680
	TACTACTCTA	TATATAATTA	CACAAATGAT	AGGACTCGTG	GCACTGCAAT	TGACAGTAAT	22740
	GATGTTGATT	GTGAACCTGT	CATAACCTAT	TCTAACATAG	GTGTTTGTA	AAATGGTGCT	22800
	TTTGTTTTTA	TTAACGTCAC	ACATTCTGAT	GGAGACGTGC	AACCAATTAG	CACTGGTAAT	22860
60	GTCACGATA	CTACAACCTT	TACCAATATCC	GTGCAAGTCG	AATATATTCA	GGTTTACTACT	22920

ES 2 170 622 B1

ACACCAGTGT CAATAGACTG TTCAAGATAT GTTTGTAAATG GTAACCCTAG GTGTAACAAA 22980  
 5 TTGTTAACAC AATACGTTTC TGCATGTCAA ACTATTGAGC AAGCACTGC AATGGGTGCC 23040  
 AGACTTGAAA ACATGGAGGT TGATTCCATG TTGTTTGTTC CTGAAAATGC CCTTAAATTG 23100  
 GCATCTGTTG AAGCATTCAA TAGTTCAGAA ACTTTAGACC CTATTTACAA AGAATGGCCT 23160  
 AATATAGGTG GTTCTTGGCT AGAAGGTCTA AAATACATAC TTCCGTCCCA TAATAGCAAA 23220  
 10 CGTAAGTATC GTTCAGCTAT AGAGGACTTG CTTTTTGATA AGGTTGTAAC ATCTGGTTTA 23280  
 GGTACAGTTG ATGAAGATTA TAAACGTTGT ACAGGTGGTT ATGACATAGC TGACTTAGTA 23340  
 TGTGCTCAAT ACTATAATGG CATCATGGTG CTACCTGGTG TGGCTAATGC TGACAAAATG 23400  
 ACTATGTACA CAGCATCCCT TGCAGGTGGT ATAACATTAG GTGCACTTGG TGGAGGCGCC 23460  
 15 GTGGCTATAC CTTTTGCAGT AGCAGTTCAG GCTAGACTTA ATTATGTTGC TCTACAAACT 23520  
 GATGTATTGA ACAAAAACCA GCAGATTCTG GCTAGTGCCT TCAATCAAGC TATTGGTAAC 23580  
 ATTACACAGT CATTGGTAA GGTAAATGAT GCTATACATC AAACATCACG AGGTCTTGCT 23640  
 ACTGTTGCTA AAGCATTGGC AAAAGTGCAA GATGTTGTCA ACATACAAGG GCAAGCTTTA 23700  
 20 AGCCACCTAA CAGTACAATT GCAAATAAT TTCCAAGCCA TTAGTAGTTC TATTAGTGAC 23760  
 ATTTATAATA GGCTTGACGA ATTGAGTGCCT GATGCACAAG TTGACAGGCT GATCACAGGA 23820  
 AGACTTACAG CACTTAATGC ATTTGTGTCT CAGACTCTAA CCAGACAAGC GGAGGTTAGG 23880  
 GCTAGTAGAC AACTTGCCAA AGACAAGGT AATGAATGCG TTAGGTCTCA GTCTCAGAGA 23940  
 25 TTCGGATTCT GTGGTAATGG TACACATTTG TTTTCACTCG CAAATGCAGC ACCAAATGGC 24000  
 ATGATTTTCT TTCACACAGT GCTATTACCA ACGGCTTATG AACTGTGAC TGCTTGCCCA 24060  
 GGTATTTGTG CTTCAGATGG TGATCGCACT TTTGGACTTG TCGTTAAAGA TGCCAGTTG 24120  
 ACTTTGTTTC GTAATCTAGA TGACAAGTTC TATTTGACCC CCAGAACTAT GTATCAGCCT 24180  
 30 AGAGTTGCAA CTAGTTCTGA CTTTGTCAA ATTGAAGGTT GCGATGTGCT GTTGTTAAT 24240  
 GCAACTGTAA GTGATTTGCC TAGTATTATA CCTGATTATA TTGATATTAA TCAGACTGTT 24300  
 CAAGACATAT TAGAAAATTT TAGACCAAAT TGGACTGTAC CTGAGTTGAC ATTTGACATT 24360  
 TTTAACGCAA CCTATTTAAA CCTGACTGGT GAAATTGATG ACTTAGAATT TAGGTCAGAA 24420  
 35 AAGCTACATA ACACCCTGT AGAACTTGCC ATTCTCATTG ACAACATTAA CAATACATTA 24480  
 GTCAATCTTG AATGGCTCAA TAGAATTGAA ACCTATGTAA AATGGCCTTG GTATGTGTGG 24540  
 CTACTAATAG GCTTAGTAGT AATATTTTGC ATACCATTAC TGCTATTTTG CTGTTGTAGT 24600  
 ACAGGTTGCT GTGGATGCAT AGGTTGTTTA GGAAGTTGTT GTCACTCTAT ATGTAGTAGA 24660  
 40 AGACAATTTG AAAATTACGA ACCAATTGAA AAAGTGCACG TCCATTAAAT TTAAAATGTT 24720  
 AATTCTATCA TCTGCTATAA TAGCAGTTGT TTCTGCTAGA GAATTTTGT AAGGATGATG 24780  
 AATAAAGTCT TTAAGAATA AACTTACGAG TCATTACAGG TCCTGTATGG ACATTGTCAA 24840  
 45 ATCCATTTAC ACATCCGTAG ATGCTGTACT TGACGAACTT GATTGTGCAT ACTTTGCTGT 24900  
 AACTCTTAAA GTAGAATTTA AGACTGGTAA ATTACTTGTG TGTATAGGTT TTGGTGACAC 24960  
 ACTTCTTGCT GCTAAGGATA AAGCATATGC TAAGCTTGGT CTCTCCATTA TTGAAGAAGT 25020  
 CAATAGTCAT ATAGTTGTTT AATATCATTAA AACACACAAA ACCCAAAGCA TTAAGTGTTA 25080  
 50 CAAAACAATT AAAGAGAGAT TATAGAAAAA CTGTCAATCT AAATCCATG CGAAAATGAT 25140  
 TGGTGGACTT TTTCTTAGTA CTCTGAGTTT TGTAAATGTT AGTAACCATT CTATTGTTAA 25200  
 TAACACAGCA AATGTGCATC ATATACAACA AGAACGTGTT ATAGTACAAC AGCATCATGT 25260  
 TGTTAGTGCT AGAACACAAA ACTATTACCC AGAGTTCAGC ATCGCTGTAC TCTTTGTATC 25320  
 55 TTTTCTAGCT TTGTACCGTA GTACAAACTT TAAGACGTGT GTCGGCATCT TAATGTTTAA 25380  
 GATTTTATCA ATGACACTTT TAGGACCTAT GCTTATAGCA TATGGTTACT ACATTGATGG 25440  
 CATTGTTACA ACAACTGTCT TATCTTTAAAG ATTTCTCTAC TTAGCATACT TTTGTTATGT 25500

60

# ES 2 170 622 B1

	TAATAGTAGG	TTTGAATTTA	TTTTATACAA	TACAACGACA	CTCATGTTTG	TACATGGCAG	25560
5	AGCTGCACCG	TTTATGAGAA	GTTCTCACAG	CTCTATTTAT	GTCACATTGT	ATGGTGGCAT	25620
	AAATTATATG	TTTGTGAATG	ACCTCACGTT	GCATTTTGTA	GACCCTATGC	TTGTAAGCAT	25680
	AGCAATACGT	GGCTTAGCTC	ATGCTGATCT	AACTGTAGTT	AGAGCAGTTG	AACTTCTCAA	25740
	TGGTGATTTT	ATTTATGTAT	TTTCACAGGA	GCCCCGATGC	GGTGTTTACA	ATGCAGCCTT	25800
10	TTCTCAGGCG	GTTCTAAACG	AAATTGACTT	AAAAGAAGAA	GAAGAAGACC	ATACCTATGA	25860
	CGTTTCCTAG	GGCATTGACT	GTATAGATG	ACAATGGAAT	GGTCATTAAC	ATCATTTTCT	25920
	GGTTCCTGTT	GATAATTATA	TTGATATTAC	TTTCAATAGC	ATTGCTAAAT	ATAATTAAGC	25980
	TATGCATGGT	GTGTTGCAAT	TTAGGAAGGA	CAGTTATTAT	TGTTCCAGCG	CAACATGCTT	26040
15	ACGATGCCTA	TAAGAATTTT	ATGCGAATTA	AAGCATAACA	CCCCGATGGA	GCACTCCTTG	26100
	CTTGAACATA	ACAAAATGAA	GATTTTGTTA	ATATTAGCGT	GTGTGATTGC	ATGCGCATGT	26160
	GGAGAACGCT	ATTGTGCTAT	GAATCCGAT	ACAGATTTGT	CATGTCGCAA	TAGTACAGCG	26220
	TCTGATTGTG	AGTCATGCTT	CAACGGAGGC	GATCTTATTT	GGCATCTTGC	AAACTGGAAC	26280
20	TTCAGCTGGT	CTATAATATT	GATCGTTTTT	ATAACTGTGC	TACAATATGG	AAGACCTCAA	26340
	TTCAGCTGGT	TCGTGTATGG	CATTAATAATG	CTTATAATGT	GGCTATTATG	GCCCCGTTGT	26400
	TTGGCTCTTA	CGATTTTTTAA	TGCATACTCG	GAATACCAAG	TGTCCAGATA	TGTAATGTTC	26460
	GGCTTTAGTA	TTGCAGGTGC	AATTGTTACA	TTTGTACTCT	GGATTATGTA	TTTTGTAAGA	26520
25	TCCATTCACT	TGTACAGAAG	GACTAAGTCT	TGGTGGTCTT	TCAACCCTGA	AACTAAAGCA	26580
	ATTCTTTGCG	TTAGTGCATT	AGGAAGAAGC	TATGTGCTTC	CTCTCGAAGG	TGTGCCAACT	26640
	GGTGTCACTC	TAACCTTGCT	TTCAGGGAAT	TTGTACGCTG	AAGGGTTCAA	AATGTCAGGT	26700
	GGTATGAACA	TCGACAATTT	ACCAAATAC	GTAATGGTTG	CATTACCTAG	CAGGACTATT	26760
30	GTCTACACAC	TTGTTGGCAA	GAAGTTGAAA	GCAAGTAGTG	CGACTGGATG	GGCTTACTAT	26820
	GTAATAATCTA	AAGCTGGTGA	TTACTCAACA	GAGGCAAGAA	CTGATAATTT	GAGTGAGCAA	26880
	CAAAAATTAT	TACATATGGT	ATAACTAAAC	TTCTAAATGG	CCAACCAGGG	ACAACGTGTC	26940
	AGTTGGGGAG	ATGAATCTAC	CAAAACACGT	GGTCGTTCCA	ATTCCCCTGG	TCGGAAGAAT	27000
35	AATAACATAC	CTCTTTCATT	CTTCAACCCC	ATAACCTCC	ACAAGGTTT	AAAATTTGG	27060
	AACTTATGTC	CGAGAGACTT	TGTACCCAAA	GGAATAGGTA	ACAGGGATCA	ACAGATTGGT	27120
	TATTGGAATA	GACAAACTCG	CTATCGCATG	GTGAAGGGCC	AACGTAAAGA	GCTTCTTGAA	27180
	AGGTGGTTCT	TCTACTACTT	AGGTACTGGA	CCTCATGCAG	ATGCCAAATT	TAAAGATAAA	27240
40	TTAGATGGAG	TTGTCTGGGT	TGCCAAGGAT	GGTGCCATGA	ACAAACCAAC	CACGCTTGGT	27300
	AGTCGTGGTG	CTAATAATGA	ATCCAAAGCT	TTGAAATTCG	ATGGTAAAGT	GCCAGGCGAA	27360
	TTTCAACTTG	AAGTTAATCA	ATCAAGAGAC	AATTCAAGGT	CACGCTCTCA	ATCTAGATCT	27420
45	CGGTCTAGAA	ATAGATCTCA	ATCTAGAGGC	AGGCAACAAT	TCAATAACAA	GAAGGATGAC	27480
	AGTGTAGAAC	AAGCTGTTCT	TGCCGCACTT	AAAAAGTTAG	GTGTTGACAC	AGAAAAACAA	27540
	CAGCAACGCT	CTCGTTCTAA	ATCTAAAGAA	CGTAGTAACT	CTAAGACAAG	AGATACTACA	27600
	CCTAAGAATG	AAAACAAACA	CACCTGGAAG	AGAACTGCAG	GTAAGGTGA	TGTGACAAGA	27660
50	TTTTATGGAG	CTAGAAGCAG	TTCAGCCAAT	TTTGGTGACA	CTGACCTCGT	TGCCAATGGG	27720
	AGCAGTGCCA	AGCATTACCC	ACAACCTGGCT	GAATGTGTTC	CATCTGTGTC	TAGCATCTG	27780
	TTTGGAAAGCT	ATTGGACTTC	AAAGGAAGAT	GGCGACCAGA	TAGAAGTCAC	GTTCCACAC	27840
	AAATACCACT	TGCCAAAGGA	TGATCCTAAG	ACTGGACAAT	TCCTTCAGCA	GATTAATGCC	27900
55	TATGCTCGTC	CATCAGAAGT	GGCAAAAGAA	CAGAGAAAAA	GAAAATCTCG	TTCTAAATCT	27960
	GCAGAAAGGT	CAGAGCAAGA	TGTGGTACCT	GATGCATTAA	TAGAAAATTA	TACAGATGTG	28020
	TTTGATGA	CACAGGTTGA	GATAATTGAT	GAGGFAACGA	ACTAAACGAG	ATGCTCGTCT	28080
60							

## ES 2 170 622 B1

5 TCCTCCATGC TGTATTTATT ACAGTTTTAA TCTTACTACT AATTGGTAGA CTCCAATTAT 28140  
TAGAAAGACT ATTACTTAAT CACTCTTTCA ATCTTAAAAC TGTCATGAC TTTAATATCT 28200  
TATATAGGAG TTTAGCAGAA ACCAGATTAC TAAAAGTGGT GCTTCGAGTA ATCTTTCTAG 28260  
TCTTACTAGG ATTTTGCTGC TACAGATTGT TAGTCACATT AATGTAAGGC AACCCGATGT 28320  
10 CTAAACTGG TTTTCCGAG GAATTACTGG TCATCGCGCT GTCTACTCTT GTACAGAATG 28380  
GTAAGCACGT GTAATAGGAG GTACAAGCAA CCCTATTGCA TATTAGGAAG TTTAGATTTG 28440  
ATTTGGCAAT GCTAGATTTA GTAATTTAGA GAAGTTTAAA GATCCGCTAC GACGAGCCAA 28500  
CAATGGAAGA GCTAACGTCT GGATCTAGTG ATTGTTTTAAA ATGTAAAATT GTTTGAAAAT 28560  
15 TTTCTTTTGG ATAGTGATAC AAAAAAAA 28588

### (2) INFORMACION DE LA SEC. ID. N°: 2:

#### 20 (i) CARACTERISTICAS DE LA SECUENCIA:

(A) LONGITUD: 21 bases

(B) TIPO: ácido nucleico

25 (C) NUMERO DE CADENAS: sencilla

(D) TOPOLOGIA: lineal

#### 30 (ii) TIPO DE MOLECULA: ácido nucleico

(A) DESCRIPCION: DNA sintético

#### 35 (xi) DESCRIPCION DE LA SECUENCIA: SEC. ID. N°: 2:

CCTAGGATTT AAATCCTAAG G

21

### (2) INFORMACION DE LA SEC. ID. N°: 3:

#### 40 (i) CARACTERISTICAS DE LA SECUENCIA:

(A) LONGITUD: 27 bases

(B) TIPO: ácido nucleico

45 (C) NUMERO DE CADENAS: sencilla

(D) TOPOLOGIA: lineal

#### 50 (ii) TIPO DE MOLECULA: ácido nucleico

(A) DESCRIPCION: DNA sintético

#### 55 (xi) DESCRIPCION DE LA SECUENCIA: SEC. ID. N°: 3:

GCGGCCGCGC CGGCGAGGCC TGTCGAC

27

### (2) INFORMACION DE LA SEC. ID. N°: 4:

#### 60 (i) CARACTERISTICAS DE LA SECUENCIA:

(A) LONGITUD: 6 bases

## ES 2 170 622 B1

	(B) TIPO: ácido nucleico	
	(C) NUMERO DE CADENAS: sencilla	
5	(D) TOPOLOGIA: lineal	
	(ii) TIPO DE MOLECULA: ácido nucleico	
10	(A) DESCRIPCION: DNA sintético	
	(xi) DESCRIPCION DE LA SECUENCIA: SEC. ID. N°: 4: GTTCGAC	6
15	(2) INFORMACION DE LA SEC. ID. N°: 5:	
	(i) CARACTERISTICAS DE LA SECUENCIA:	
20	(A) LONGITUD: 22 bases	
	(B) TIPO: ácido nucleico	
	(C) NUMERO DE CADENAS: sencilla	
25	(D) TOPOLOGIA: lineal	
	(ii) TIPO DE MOLECULA: ácido nucleico	
30	(A) DESCRIPCION: DNA sintético	
	(xi) DESCRIPCION DE LA SECUENCIA: SEC. ID. N°: 5: GCTAGCCCAG GCGCGCGGTA CC	22
35		
40		
45		
50		
55		
60		



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA

① ES 2 170 622

② Nº de solicitud: 009902673

③ Fecha de presentación de la solicitud: 03.12.1999

④ Fecha de prioridad:

## INFORME SOBRE EL ESTADO DE LA TÉCNICA

⑤ Int. Cl.7: C12N 7/04, 15/50, 15/86, A61K 39/12, 39/295

### DOCUMENTOS RELEVANTES

Categoría	Documentos citados	Reivindicaciones afectadas
X	WO 9943842 A (NEUROVIR INC) 02.09.1999, página 2, línea 11 - página 6, línea 17.	1-16
A	WO 9906582 A (KOSZINOWSKI U.) 11.02.1999, todo el documento.	1-16
A	WO 9503400 A (JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE) 02.02.1995, todo el documento.	1-16
A	WO 9734008 A (CYANAMID IBÉRICA, S.A.) 18.09.1997, todo el documento.	1-16

#### Categoría de los documentos citados

X: de particular relevancia

Y: de particular relevancia combinado con otro/s de la misma categoría

A: refleja el estado de la técnica

O: referido a divulgación no escrita

P: publicado entre la fecha de prioridad y la de presentación de la solicitud

E: documento anterior, pero publicado después de la fecha de presentación de la solicitud

#### El presente informe ha sido realizado

para todas las reivindicaciones

para las reivindicaciones nº:

Fecha de realización del informe

02.07.2002

Examinador

A. Collados Martín Posadillo

Página

1/1