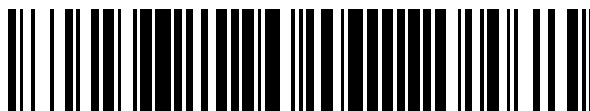


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 396 127**

51 Int. Cl.:

C12N 7/00 (2006.01)

C07K 14/165 (2006.01)

G01N 33/569 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **02.12.2004 E 04805624 (6)**

97 Fecha y número de publicación de la concesión europea: **19.09.2012 EP 1697507**

54 Título: **Uso de proteínas y peptidos codificados por el genoma de una nueva cepa de coronavirus asociado al SRAS**

30 Prioridad:

02.12.2003 FR 0314152

02.12.2003 FR 0314151

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

19.02.2013

73 Titular/es:

**INSTITUT PASTEUR (33.3%)
25-28, RUE DU DOCTEUR ROUX
75724 PARIS CEDEX 15, FR;
CENTRE NATIONAL DE LA RECHERCHE
SCIENTIFIQUE (CNRS) (33.3%) y
UNIVERSITE PARIS VII (33.3%)**

72 Inventor/es:

**VAN DER WERF, SYLVIE;
ESCRIOU, NICOLAS;
CRESCENZO-CHAIGNE, BERNADETTE;
MANUGUERRA, JEAN-CLAUDE;
KUNST, FREDERIK, INSTITUT PASTEUR;
CALLENDRET, BENOÎT;
BETTON, JEAN-MICHEL;
LORIN, VALÉRIE;
GERBAUD, SYLVIE;
BURGUIERE, ANA MARIA;
AZEBI, SALIHA;
CHARNEAU, PIERRE;
TANGY, FRÉDÉRIC;
COMBREDT, CHANTAL;
DELAGNEAU, JEAN-FRANÇOIS y
MARTIN, MONIQUE**

74 Agente/Representante:

DURÁN MOYA, Luis Alfonso

ES 2 396 127 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

DESCRIPCIÓN

Uso de proteínas y péptidos codificados por el genoma de una nueva cepa de coronavirus asociado al SRAS

La presente invención se refiere a una nueva cepa de coronavirus asociada al síndrome respiratorio agudo severo (SRAS), procedente de una extracción catalogada bajo el nº 031589 y extraída en Hanoi (Vietnam), a moléculas de ácido nucleico procedentes de su genoma, a las proteínas y péptidos codificados por dichas moléculas de ácido nucleico, así como a sus aplicaciones, en particular como agentes reactivos de diagnóstico y/o como vacuna.

El coronavirus es un virus de ARN monocatenario, de polaridad positiva, de aproximadamente 30 kilobases, que se replica en el citoplasma de células hospedantes; el extremo 5' del genoma tiene una estructura en capuchón y el extremo 3' comprende una cola polyA. Este virus está envuelto y comprende, en su superficie, estructuras peploméricas denominadas espículas.

El genoma comprende los cuadros abiertos de lectura, u ORF, siguientes, desde su extremo 5' hacia su extremo 3'; ORF1a y ORF1b, que corresponden a las proteínas del complejo de transcripción-replicación, y ORF-S, ORF-E, ORF-M y ORF-N, que corresponden a las proteínas estructurales S, E, M y N. Comprende también unos ORF que corresponden a proteínas de función desconocida codificadas por: la región situada entre el ORF-S y el ORF-E y que se superpone a esta última, la región situada entre el ORF-M y el ORF-N, y la región incluida entre el ORF-N.

La proteína S es una glicoproteína membranaria (200-220 kDa), que se presenta en forma de espículas o "pinchos" que emergen de la superficie de envoltura viral. Es responsable de la adhesión del virus a los receptores de la célula hospedante, y de la inducción de la fusión de la envoltura viral con la membrana celular.

La pequeña proteína de envoltura (E) también denominada sM (*small membrane*), que es una proteína transmembranaria no glicosilada de aproximadamente 10 kDa, es la proteína presente en cantidad más baja en el virión. Desempeña un papel motor en el proceso del brote de coronavirus que se produce a nivel del compartimiento intermedio en el retículo endoplásmico y el aparato de Golgi.

La proteína M, o proteína matriz (25-30 kDa), es una glicoproteína membranaria más abundante que está integrada en la partícula viral por una interacción M/E, mientras que la incorporación de S en las partículas está dirigida por una interacción S/M. Parece ser importante para la maduración viral de los coronavirus y para la determinación del sitio al nivel del cual las partículas virales se ensamblan.

La proteína N o proteína de nucleocápside (45-50 kDa), que es la más conservada entre las proteínas estructurales de los coronavirus, es necesaria para encapsidar el ARN genómico, y después para dirigir su incorporación en el virión. Esta proteína está probablemente también implicada en la replicación del ARN.

Cuando una célula hospedante es infectada, el cuadro de lectura (ORF) situado en 5' del genoma viral se traduce en una poliproteína que es escindida por las proteasas virales, y libera entonces varias proteínas no estructurales, tales como ARN-polimerasa, ARN dependiente (Rep) ATPasa helicasa (Hel). Estas dos proteínas están implicadas en la replicación del genoma viral, así como en la generación de transcritos que son utilizados en la síntesis de las proteínas virales. Los mecanismos por los cuales se producen estos ARNm sub-genómicos no están totalmente comprendidos; sin embargo, hechos recientes indican que las secuencias de regulación de la transcripción en el extremo 5' de cada gen representan señales que regulan la transcripción discontinua de los ARNm sub-genómicos.

Las proteínas de la membrana viral (proteínas S, E y M) están insertadas en el compartimiento intermedio, mientras que el ARN replicado (hebra +) se ensambla con la proteína N (nucleocápside). Este complejo proteína-ARN se asocia después con la proteína M incluida en las membranas del retículo endoplásmico y las partículas virales se forman cuando el complejo de la nucleocápside brota en el retículo endoplásmico. El virus migra después a través del complejo del Golgi y eventualmente sale de la célula, por ejemplo mediante exocitosis. El sitio de adhesión del virus a la célula hospedante se encuentra al nivel de la proteína S.

Los coronavirus son responsables del 15 al 30% de los catarros en el ser humano, y de infecciones respiratorias o digestivas en los animales, en particular gatos (FIPV: *Feline infectious peritonitis virus*), aves de corral (IBV: *Avian Infectious bronchitis virus*), ratones (MHV: *Mouse Hepatitis virus*), cerdos (TGEV: *Transmissible gastroenteritis virus*, PEDV: *Porcine Epidemic Diarrhea virus*, PRCov: *Porcine Respiratory Coronavirus*, HEV: *Hemagglutinating encephalomyelitis Virus*) y los bovinos (Bcov: *Bovine coronavirus*).

En general, cada coronavirus afecta sólo a una especie; en los individuos inmunocompetentes, la infección induce unos anticuerpos eventualmente neutralizantes, y una inmunidad celular, capaces de destruir las células infectadas.

Una epidemia de neumonía atípica, denominada síndrome respiratorio agudo severo (SARS o *Severe acute respiratory syndrome*, SRAS en español), se ha propagado en diferentes países (Vietnam, Hong-Kong, Singapur, Tailandia y Canadá) durante el primer trimestre de 2003, a partir de un foco inicial aparecido en China, en el último trimestre de 2002. La severidad de esta enfermedad es tal, que su índice de mortalidad es aproximadamente del 3 al

6%. La determinación del agente causante de esta enfermedad se ha investigado por numerosos laboratorios, en todo el mundo.

En marzo de 2003, se ha aislado un nuevo coronavirus (SARS-CoV, SARS virus o virus SRAS en español), en asociación con unos casos de síndrome respiratorio agudo severo (T.G.KSIAZEK y otros, The New England Journal of Medicine, 2003, 348, 1319-1330; C. DROSTEN y otros, The New England Journal of Medicine, 2003, 348, 1967-1976; Peiris y otros, Lancet, 2003, 361, 1319-).

Se obtuvieron así unas secuencias genómicas de este nuevo coronavirus, en particular las del aislado Urbani (Genbank nº de acceso AY274119.3 y A. MARRA y otros, Science, 1 de mayo, 2003, 300, 1399-1404) y del aislado de Toronto (Tor2, Genbank nº de acceso AY 278741 y A. ROTA y otros, Science, 2003, 300, 1394-1399).

La organización del genoma es comparable a la de los demás coronavirus conocidos, permitiendo así confirmar la pertenencia del SRAS-CoV a la familia de los *Coronaviridae*; se han identificado en particular los cuadros de lectura abiertos ORF1a y 1b, y los cuadros de lectura abiertos que corresponden a las proteínas S, E, M y N, así como a proteínas codificadas por: la región situada entre ORF-S y ORF-E (ORF3), la región situada entre ORF-S y ORF-E y que se superpone a ORF-E (ORF4), la región situada entre ORF-M y ORF-N (ORF7 a ORF11) y la región que corresponde a ORF-N (ORF13 y ORF14).

Se han puesto en evidencia siete diferencias entre las secuencias de los aislados Tor2 y Urbani; 3 corresponden a mutaciones silenciosas (c/t en la posición 16622 y a/g en la posición 19064 de ORF1b, t/c en la posición 24872 de ORF-S), y 4 modifican la secuencia en aminoácidos respectivamente: las proteínas codificadas por ORF1a (c/t en la posición 7919 que corresponde a la mutación A/V), la proteína S (g/t en la posición 23220 que corresponde a la mutación A/S), la proteína codificada por ORF3 (a/g en la posición 25298 que corresponde la mutación R/G) y de la proteína M (t/c en la posición 26857 que corresponde a la mutación S/P).

Además, el análisis filogenético muestra que el SRAS-CoV está distante de los demás coronavirus, y que no ha aparecido ni por mutación de coronavirus respiratorios humanos, ni por recombinación entre unos coronavirus conocidos (para una revisión, véase Homes J.C.I., 2003, 111, 1605-1609).

La puesta en evidencia y la consideración de nuevas variantes son importantes para la realización de agentes reactivos de detección y de diagnóstico de SRAS suficientemente sensibles y específicos, así como composiciones inmunógenas aptas para proteger las poblaciones contra unas epidemias de SRAS.

Los inventores han puesto en evidencia ahora otra cepa de coronavirus asociado al SRAS, que se distingue de los aislados Tor2 y Urbani.

La presente invención es tal como se define por las reivindicaciones 1 a 22 más adelante.

Los inventores describen una cepa aislada o purificada de coronavirus humano asociado al síndrome respiratorio agudo severo, caracterizada porque su genoma presenta, en forma de ADN complementario, un codón de serina en la posición 23220-23222 del gen de la proteína S, o un codón de glicina en la posición 25298-25300 del gen de ORF3, y un codón de alanina en la posición 7918-7920 de ORF1a, o un codón de serina en la posición 26857-26859 del gen de la proteína M, siendo dichas posiciones indicadas en referencia a la secuencia Genbank AY274119.3

Según otro modo de realización ventajoso de dicha cepa, el equivalente ADN de su genoma presenta una secuencia que corresponde a la secuencia SEC ID nº 1; esta cepa de coronavirus procede de la extracción de un lavado broncoalveolar de un paciente afectado de SRAS, catalogado bajo el nº 031589 y efectuado en el hospital francés de Hanoi (Vietnam).

Dicha secuencia SEC ID nº 1 es la del ácido desoxirribonucleico, que corresponde a la molécula de ácido ribonucleico del genoma de la cepa aislada de coronavirus, tal como se definió anteriormente.

La secuencia SEC ID nº 1 se distingue de la secuencia Genbank AY274119.3 (aislado Tor2) porque posee las mutaciones siguientes:

- g/t en la posición 23220; el codón de alanina (gct) en la posición 577 de la secuencia en aminoácidos de la proteína S de Tor2 está sustituido con un codón de serina (tct),
- a/g en la posición 25298; el codón de arginina (aga) en la posición 11 de la secuencia de aminoácidos de la proteína codificada por ORF3 de Tor2 está sustituido con un codón de glicina (gga).

Además, la secuencia SEC ID nº 1 se distingue de la secuencia Genbank AY278741 (aislado Urbani) porque posee las mutaciones siguientes:

- t/c en la posición 7919; el codón de valina (gtt) en la posición 2552 de la secuencia en aminoácidos de la

proteína codificada por ORF1a está sustituido con un codón de alanina (gct),

- t/c en la posición 16622: esta mutación no modifica la secuencia en aminoácidos de las proteínas codificadas por ORF1b (mutación silenciosa),
- g/a en la posición 19064: esta mutación no modifica la secuencia de aminoácidos de las proteínas codificadas por ORF1b (mutación silenciosa),
- c/t en la posición 24872: esta mutación no modifica la secuencia de aminoácidos de la proteína S, y
- -c/t en la posición 26857: el codón de prolina (ccc) en la posición 154 de la secuencia de aminoácidos de la proteína M está sustituido con un codón de serina (tcc).

En ausencia de mención particular, las posiciones de las secuencias nucleotídicas y peptídicas están indicadas con referencia a la secuencia Genbank AY274119.3.

Los inventores describen también un polinucleótido aislado o purificado, caracterizado porque su secuencia es la del genoma de la cepa aislada de coronavirus, tal como se ha definido anteriormente.

Según un modo de realización ventajoso de dicho polinucleótido, este presenta la secuencia SEC ID nº 1.

Los inventores describen también un polinucleótido aislado o purificado, caracterizado porque su secuencia híbrida en condiciones de fuerte rigor, con la secuencia del polinucleótido, tal como se ha definido anteriormente.

Los términos "aislado o purificado" significan modificado "por la mano del ser humano" a partir del estado natural; dicho de otra manera, si un objeto existe en la naturaleza, este se denomina aislado, o purificado si se ha modificado o extraído de su entorno natural, o las dos. Por ejemplo, un polinucleótido, o una proteína/péptido naturalmente presente en un organismo vivo, no es ni aislado, ni purificado; en cambio, el mismo polinucleótido o proteína/péptido separado de las moléculas que coexisten en su entorno natural, obtenido por clonación, amplificación y/o síntesis química, está aislado en el sentido de la presente invención. Además, un polinucleótido o una proteína/péptido introducido en un organismo por transformación, manipulación genética o cualquier otro método, está "aislado" incluso si está presente en dicho organismo. El término purificado, tal como se utiliza en la presente invención, significa que las proteínas/péptidos, según la invención, están esencialmente libres de asociación con otras proteínas o polipéptidos, como lo es, por ejemplo, el producto purificado del cultivo de células hospedantes recombinantes, o el producto purificado a partir de una fuente no recombinante.

Se entiende por condiciones de hibridación de fuerte rigor, condiciones de temperatura y de fuerza iónica seleccionadas de tal manera que permiten el mantenimiento de la hibridación específica y selectiva entre los polinucleótidos complementarios.

A título ilustrativo, condiciones de fuerte rigor con el fin de definir los polinucleótidos anteriores, son ventajosamente las siguientes: la hibridación ADN-ADN o ADN-ARN se realiza en dos etapas: (1) prehibridación a 42°C durante 3 horas en tampón de fosfato (20 mM pH 7,5), que contiene 5 x SSC (1 x SSC corresponde a una solución de 0,15 M de NaCl + 0,015 M de citrato de sodio), el 50% de formamida, el 7% de dodecilsulfato de sodio (SDS), 10 x Denhardt's, 5% de sulfato de dextrano y el 1% de ADN de esperma de salmón; (2) hibridación durante 20 horas a 42°C seguida de 2 lavados de 20 minutos a 20°C en 2 x SSC + el 2% SDS, 1 lavado de 20 minutos a 20°C en 0,1 x SSC + el 0,1% SDS. El último lavado se realiza en 0,1 x SSC + el 0,1% SDS durante 30 minutos a 60°C.

Los inventores describen asimismo un fragmento representativo del polinucleótido, tal como se ha definido anteriormente, caracterizado porque es susceptible de ser obtenido o bien mediante el uso de enzimas de restricción, cuyos sitios de reconocimiento y de escisión están presentes en dicho polinucleótido, tal como se ha definido anteriormente, o bien mediante amplificación con la ayuda de cebadores oligonucleotídicos específicos de dicho polinucleótido, tal como se ha definido anteriormente, o bien mediante transcripción *in vitro*, o bien mediante síntesis química.

Según un modo de realización ventajoso de dicho fragmento, éste se selecciona del grupo constituido por: el ADNc, que corresponde a al menos un cuadro de lectura abierto (ORF), seleccionado entre: ORF1a, ORF1b, ORF-S, ORF-E, ORF-M, ORF-N, ORF3, ORF4, ORF7 a ORF11, ORF13 y ORF 14, y el ADNc, que corresponde a los extremos 5' o 3' no codificantes de dicho polinucleótido.

Según una disposición ventajosa de este modo de realización, dicho fragmento presenta una secuencia seleccionada del grupo constituido por:

- las secuencias SEC ID nº 2 y 4, que representan el ADNc que corresponde al ORF-S, que codifica la proteína S,
- las secuencias SEC ID nº 13 y 15, que representan el ADNc que corresponde al ORF-E, que codifica la proteína

E,

- las secuencias SEC ID nº 16 y 18, que representan el ADNc que corresponde al ORF-M, que codifica la proteína M,
- las secuencias SEC ID nº 36 y 38, que representan el ADNc que corresponde al ORF-N, que codifica la proteína N,
- las secuencias que representan los ADNc que corresponden respectivamente: a ORF1a y ORF1b (ORF1ab, SEC ID nº 31), a los ORF3 y ORF4 (SEC ID nº 7, 8), a los ORF 7 a 11 (SEC ID nº 19, 20), al ORF13 (SEC ID nº 32) y al ORF14 (SEC ID nº 34), y
- las secuencias que representan los ADNc que corresponden respectivamente a los extremos 5' (SEC ID nº 39 y 72) y 3' no codificantes (SEC ID nº 40, 73) de dicho polinucleótido.

Los inventores describen también un fragmento del ADNc que codifica la proteína S, tal como se ha definido anteriormente, caracterizado porque presenta una secuencia seleccionada del grupo constituido por las secuencias SEC ID nº 5 y 6 (fragmentos Sa y Sb).

Los inventores describen también un fragmento del ADNc que corresponde a los ORF1a y ORF1b, tal como se han definido anteriormente, caracterizado porque presenta una secuencia seleccionada del grupo constituido por las secuencias SEC ID nº 41 a 54 (fragmentos L0 a L12).

Los inventores describen también un fragmento del polinucleótido, tal como se ha definido anteriormente, caracterizado porque presenta al menos 15 bases o pares de bases consecutivas de la secuencia del genoma de dicha cepa, que incluye al menos una de las situadas en las posiciones 7979, 16622, 19064, 23220, 24872, 25298 y 26857. Preferiblemente, se trata de un fragmento de 20 a 2500 bases o pares de bases, de manera preferente de 20 a 400.

Según un modo de realización ventajoso de dicho fragmento, este incluye al menos una pareja de bases o de pares de bases que corresponden a las posiciones siguientes: 7919 y 23220, 7919 y 25298, 16622 y 23220, 19064 y 23220, 16622 y 25298, 19064 y 25298, 23220 y 24872, 23220 y 26857, 24872 y 25298, 25298 y 26857.

Los inventores describen también unos cebadores de al menos 18 bases, aptos para amplificar un fragmento del genoma de un coronavirus asociado al SRAS, o del equivalente ADN de este.

Según un modo de realización de dichos cebadores, estos se seleccionan del grupo constituido por:

- el par de cebadores nº 1, que corresponde respectivamente a las posiciones 28507 a 28522 (cebador sentido, SEC ID nº 60) y 28774 a 28759 (cebador anti-sentido, SEC ID nº 61) de la secuencia del polinucleótido, tal como se ha definido anteriormente,
- el par de cebadores nº 2, que corresponde respectivamente a las posiciones 28375 a 28390 (cebador sentido, SEC ID nº 62) y 28702 a 28687 (cebador anti-sentido, SEC ID nº 63) de la secuencia del polinucleótido, tal como se ha definido anteriormente, y
- el par de cebadores constituido de los cebadores SEC ID nº 55 y 56.

Los inventores describen también una sonda para detectar la presencia del genoma de un coronavirus asociado al SRAS o de un fragmento de este, caracterizada porque se selecciona del grupo constituido por: los fragmentos tales como se han definido anteriormente, y los fragmentos que corresponden a las posiciones siguientes de la secuencia del polinucleótido, tal como se ha definido anteriormente: 28561 a 28586, 28588 a 28608, 28541 a 28563 y 28565 a 28589 (SEC ID nº 64 a 67).

Las sondas y cebadores pueden ser marcados directa o indirectamente por un compuesto radioactivo o no radioactivo, mediante métodos bien conocidos por el experto en la materia, a fin de obtener una señal detectable y/o cuantificable. Entre los isótopos radioactivos utilizados, se pueden citar ^{32}P , ^{33}P , ^{35}S , ^3H ó ^{125}I . La entidades no radioactivas se seleccionan entre los ligantes, tales como la biotina, la avidina, la estreptavidina, la digoxigenina, los haptenos, los colorantes, los agentes luminescentes tales como los agentes radioluminescentes, quemoluminescentes, bioluminescentes, fluorescentes, fosforescentes.

Las sondas y cebadores engloban las sondas y los cebadores marcados derivados de las secuencias anteriores.

Tales sondas y cebadores son útiles para el diagnóstico de la infección por un coronavirus asociado al SRAS.

Los inventores describen también un método de detección de un coronavirus asociado al SRAS, a partir de una

muestra biológica, método el cual se caracteriza porque comprende al menos:

- (a) extraer los ácidos nucleicos presentes en dicha muestra biológica,
- (b) amplificar un fragmento del ORF-N mediante RT-PCR con la ayuda de un par de cebadores tales como se han definido anteriormente, y
- (c) detectar mediante cualquier medio apropiado de los productos de amplificaciones obtenidos en (b).

Los productos de amplificaciones (amplicones) en (b) son de 268 pb para el par de cebadores nº 1 y de 328 pb para el par de cebadores nº 2.

Según un modo de aplicación ventajoso de dicho procedimiento, la etapa (b) de detección se realiza con la ayuda de al menos una sonda que corresponde a las posiciones 28561 a 28586, 28588 a 28608, 28541 a 28563 y 28565 a 28589 de la secuencia de polinucleótido, tal como se definió anteriormente.

Preferentemente, el genoma de coronavirus asociado al SRAS se detecta, y eventualmente se cuantifica por PCR en tiempo real, con la ayuda del par de cebadores nº 2 y de las sondas que corresponden a las posiciones 28541 a 28563, y 28565 a 28589, marcadas con unos compuestos diferentes, especialmente unos agentes fluorescentes diferentes.

LA RT-PCR en tiempo real que aplica este par de cebadores y esta sonda es muy sensible, ya que permite detectar 10^2 copias de ARN y hasta 10 copias de ARN, es además fiable y reproducible.

La presente solicitud engloba los polidesoxirribonucleótidos y los polirribonucleótidos monocatenarios, bicatenarios y tricatenarios, que corresponden a la secuencia del genoma de la cepa aislada de coronavirus y de sus fragmentos, tales como se han definido anteriormente, así como a sus secuencias complementarias, sentido o anti-sentido, en particular los ARN y los ADNc que corresponden a la secuencia del genoma y de sus fragmentos, tales como se han definido anteriormente.

Los inventores describen también los fragmentos de amplificación obtenidos con la ayuda de cebadores específicos del genoma de la cepa purificada o aislada, tal como se ha definido anteriormente, en particular con la ayuda de cebadores y de pares de cebadores, tales como se han definido anteriormente, los fragmentos de restricción constituidos por, o que comprenden la secuencia de los fragmentos, tales como se han definido anteriormente, los fragmentos obtenidos por transcripción *in vitro* a partir de un vector que contiene la secuencia SEC ID nº 1 o un fragmento tal como se ha definido anteriormente, así como fragmentos obtenidos por síntesis química. Ejemplos de fragmentos de restricción se deducen del mapa de restricción de la secuencia SEC ID nº 1, ilustrado por la figura 13. Dichos fragmentos están o bien en forma de fragmentos aislados, o bien en forma de mezclas de fragmentos. La presente solicitud engloba también los fragmentos modificados, con respecto a los anteriores, por escisión, o adición de nucleótidos en una proporción de aproximadamente el 15% con respecto a la longitud de los fragmentos anteriores, y/o modificados al nivel de la naturaleza de los nucleótidos, a partir del momento en el que los fragmentos nucleotídicos modificados conserven una capacidad de hibridación con las secuencias de ARN genómico o antígeno del aislado, tal como se ha definido anteriormente.

Las moléculas de ácido nucleico, según la invención, se obtienen mediante los métodos clásicos, conocidos en sí, siguiendo los protocolos estándares, tales como los descritos en *Current Protocols in Molecular Biology* (Frederick M. AUSUBEL, 2000, Wiley and son Inc, Library of Congress, USA). Por ejemplo, se pueden obtener mediante amplificación de una secuencia nucleica por PCR o RT-PCR, o bien mediante síntesis química total o parcial.

Los inventores describen también un chip o filtro de ADN o de ARN, caracterizado porque comprende al menos un polinucleótido o uno de sus fragmentos, tales como se han definido anteriormente.

Los chip o filtros de ADN o de ARN se preparan mediante los métodos clásicos, conocidos en sí, como por ejemplo injerto químico o electroquímico de oligonucleótido sobre soporte de vidrio o de nylon.

La presente invención tiene también por objeto un vector de clonación y/o de expresión recombinante, en particular un plásmido, un virus, un vector viral o un fago, que comprende un fragmento de ácido nucleico, según la presente invención. Preferiblemente, dicho vector recombinante es un vector de expresión en el que dicho fragmento de ácido nucleico está colocado bajo el control de elementos reguladores de la transcripción y de la traducción apropiados. Además, dicho vector puede comprender unas secuencias (etiquetas o *tag*) fusionadas en fase con el extremo 5' y/o 3' de dicho inserto, útiles para la inmovilización y/o la detección y/o la purificación de la proteína expresada a partir de dicho vector.

Estos vectores están contruidos e introducidos en células hospedantes mediante métodos clásicos de ADN recombinante y de ingeniería genética, que son conocidos en sí. Numerosos vectores en los que se pueden insertar una molécula de ácido nucleico de interés, a fin de introducirlo y mantenerlo en una célula hospedante, son

conocidos en sí; la elección de un vector apropiado depende del uso considerado para este vector (por ejemplo, replicación de la secuencia de interés, expresión de esta secuencia, mantenimiento de la secuencia en forma extracromosómica o bien integración en el material cromosómico del hospedante), así como de la naturaleza de la célula hospedante.

Los inventores describen los plásmidos siguientes:

- el plásmido denominado SRAS-S, comprendido en la cepa bacteriana depositada bajo el n° 1-3059, el 20 de junio de 2003, en la Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15; contiene la secuencia de ADNc que para codifica la proteína S de la cepa de SRAS-CoV procedente de la extracción catalogada bajo el n° 031589, secuencia que corresponde a los nucleótidos de las posiciones 21406 a 25348 (SEC ID n° 4), en referencia a la secuencia Genbank AY274119.3,
- el plásmido denominado SRAS-S1, comprendido en la cepa bacteriana depositada bajo el n° I-3020, el 12 de mayo de 2003, en la Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15; contiene un fragmento 5' de la secuencia de ADNc que codifica la proteína S de la cepa de SRAS-CoV procedente de la extracción catalogada bajo el n° 031589, tal como se ha definido anteriormente, fragmento que corresponde a los nucleótidos de las posiciones 21406 a 23454 (SEC ID n° 5), en referencia a la secuencia Genbank AY274119.3 Tor2,
- el plásmido denominado SRAS-S2, comprendido en la cepa bacteriana depositada bajo el n° I-3019, el 12 de mayo de 2003, en la Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15; contiene un fragmento 3' de la secuencia de ADNc que codifica la proteína S de la cepa de SRAS-CoV procedente de la extracción catalogada bajo el n° 031589, tal como se ha definido anteriormente, fragmento que corresponde a los nucleótidos de las posiciones 23322 a 25348 (SEC ID n° 6), en referencia a la secuencia Genbank n° de acceso AY274119.3,
- el plásmido denominado SRAS-SE, comprendido en la cepa bacteriana depositada bajo el n° 1-3126, el 13 de noviembre de 2003, en la Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15; contiene el ADNc que corresponde a la región situada entre ORF-S y ORF-E y que se superpone al ORF-E de la cepa de SRAS-CoV procedente de la extracción catalogada bajo el n° 031589, tal como se ha definido anteriormente, región que corresponde a los nucleótidos de las posiciones 25110 a 26244 (SEC ID n° 8), en referencia a la secuencia Genbank n° de acceso AY274119.3,
- el plásmido denominado SRAS-E, comprendido en la cepa bacteriana depositada bajo el n° 1-3046, el 28 de mayo de 2003, en la Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15; contiene la secuencia de ADNc que codifica la proteína E de la cepa de SRAS-CoV procedente de la extracción catalogada bajo el n° 031589, tal como se ha definido anteriormente, secuencia que corresponde a los nucleótidos de las posiciones 26082 a 26413 (SEC ID n° 15), en referencia a la secuencia Genbank n° de acceso AY274119.3,
- el plásmido denominado SRAS-M; comprendido en la cepa bacteriana depositada bajo el n° 1-3047, el 28 de mayo de 2003, en la Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15; contiene la secuencia de ADNc que codifica la proteína M de la cepa de SRAS-CoV procedente de la extracción catalogada bajo el n° 031589, tal como se ha definido anteriormente; secuencia que corresponde a los nucleótidos de las posiciones 26330 a 27098 (SEC ID N° 18), en referencia a la secuencia Genbank n° de acceso AY274119.3,
- el plásmido denominado SRAS-MN, comprendido en la cepa bacteriana depositada bajo el n° I-3125, el 13 de noviembre de 2003, en la Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15; contiene la secuencia de ADNc que corresponde a la región situada entre el ORF-M y el ORF-N de la cepa de SRAS-CoV, procedente de la extracción catalogada bajo el n° 031589 y extraída en Hanoi, tal como se ha definido anteriormente, secuencia que corresponde a los nucleótidos de las posiciones 26977 a 28218 (SEC ID n° 20), en referencia a la secuencia Genbank n° de acceso AY274119.3,
- el plásmido denominado SRAS-N, comprendido en la cepa bacteriana depositada bajo el n° 1-3048, el 5 de junio de 2003, en la Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15; contiene el ADNc que codifica la proteína N de la cepa de SRAS-CoV procedente de la extracción catalogada bajo el n° 031589, tal como se ha definido anteriormente, secuencia que corresponde a los nucleótidos de las posiciones 28054 a 29430 (SEC ID N° 38), en referencia a la secuencia Genbank n° de acceso AY274119.3; este plásmido comprende así un inserto de secuencia SEC ID n° 38, y está comprendido en una cepa bacteriana que se ha depositado bajo el n° 1-3048, el 5 de junio de 2003, en la Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15,
- el plásmido denominado SRAS-5'NC, comprendido en la cepa bacteriana depositada bajo el n° I-3124, el 7 de noviembre de 2003, en la Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15; contiene el ADNc que corresponde al extremo 5' no codificante del genoma de la cepa de SRAS-

CoV procedente de la extracción catalogada bajo el n° 031589, tal como se ha definido anteriormente, secuencia que corresponde a los nucleótidos de las posiciones 1 a 204 (SEC ID n° 39), en referencia a la secuencia Genbank n° de acceso AY274119.3,

- el plásmido denominado SRAS-3'NC, comprendido en la cepa bacteriana depositada bajo el n° 1-3123 el 7 de noviembre de 2003, en la Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15; contiene el ADNc que corresponde al extremo 3' no codificante del genoma de la cepa de SRAS-CoV procedente de la extracción catalogada bajo el n° 031589, tal como se ha definido anteriormente, secuencia que corresponde a la situada entre el nucleótido en la posición 28933 a 29727 (SEC ID n° 40), en referencia a la secuencia Genbank n° de acceso AY274119.3, se termina por una serie de nucleótidos a.,

- el plásmido de expresión denominado pIV2.3N, que contiene un fragmento de ADNc que codifica una fusión C-terminal de la proteína N (SEC ID n° 37) con una etiqueta de polihistidina,

- el plásmido de expresión denominado pIV2.3S_C, que contiene un fragmento de ADNc que codifica una fusión C-terminal del fragmento que corresponde a las posiciones 475 a 1193 de la secuencia en aminoácidos de la proteína S (SEC ID n° 3) con una etiqueta de polihistidina,

- el plásmido de expresión pIV2.3S_L, que contiene un fragmento de ADNc que codifica una fusión C-terminal del fragmento que corresponde a las posiciones 14 a 1193 de la secuencia en aminoácidos de la proteína S (SEC ID n° 3) con una etiqueta de polihistidina,

- el plásmido de expresión denominado pIV2.4N, que contiene un fragmento de ADNc que codifica una fusión N-terminal de la proteína N (SEC ID n° 3) con una etiqueta de polihistidina,

- el plásmido de expresión denominado pIV2.4S_C o pIV2.4S₁, que contiene un inserto que codifica una fusión N-terminal del fragmento que corresponde a las posiciones 475 a 1193 de la secuencia en aminoácidos de la proteína S (SEC ID n° 3) con una etiqueta de polihistidina, y

- el plásmido de expresión denominado pIV2.4S_L que contiene un fragmento de ADNc que codifica una fusión N-terminal del fragmento que corresponde a las posiciones 14 a 1193 de la secuencia en aminoácidos de la proteína S (SEC ID n° 3) con una etiqueta de polihistidina.

Según una disposición ventajosa del plásmido de expresión, tal como se ha definido anteriormente, está comprendido en una cepa bacteriana depositada bajo el n° I-3117, el 23 de octubre de 2003, en la Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15.

Según otra disposición ventajosa del plásmido de expresión, tal como se ha definido anteriormente, está comprendido en una cepa bacteriana depositada bajo el n° I-3118, el 23 de octubre de 2003, en la Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15.

Según otra disposición del plásmido de expresión, tal como se ha definido anteriormente, está comprendido en una cepa bacteriana depositada en la CNCM, 25 rue du Docteur Roux, 75724 Paris Cedex 15 bajo los números siguientes:

- a) cepa n° I-3118, depositada el 23 de octubre de 2003,
- b) cepa n° I-3019, depositada el 12 de mayo de 2003,
- c) cepa n° I-3020, depositada el 12 de mayo de 2003,
- d) cepa n° I-3059, depositada el 20 de junio de 2003,
- e) cepa n° I-3323, depositada el 22 de noviembre de 2004,
- f) cepa n° I-3324, depositada el 22 de noviembre de 2004,
- g) cepa n° I-3326, depositada el 1 de diciembre de 2004,
- h) cepa n° I-3327, depositada el 1 de diciembre de 2004,
- i) cepa n° I-3332, depositada el 1 de diciembre de 2004,
- j) cepa n° I-3333, depositada el 1 de diciembre de 2004,
- k) cepa n° I-3334, depositada el 1 de diciembre de 2004,

- l) cepa n° I-3335, depositada el 1 de diciembre de 2004,
- m) cepa n° I-3336, depositada el 1 de diciembre de 2004,
- n) cepa n° I-3337, depositada el 1 de diciembre de 2004,
- o) cepa n° I-3338, depositada el 2 de diciembre de 2004,
- p) cepa n° I-3339, depositada el 2 de diciembre de 2004,
- q) cepa n° I-3340, depositada el 2 de diciembre de 2004,
- r) cepa n° I-3341, depositada el 2 de diciembre de 2004.

Los inventores describen también un inserto de ácido nucleico de origen viral, caracterizado porque está contenido en cualquiera de las cepas tales como se definieron anteriormente en a)-r).

Los inventores describen también un ácido nucleico que comprende un gen sintético que permite una expresión optimizada de la proteína S en unas células eucariotas, caracterizado porque posee la secuencia SEC ID n° 140.

Los inventores describen también un vector de expresión que comprende un ácido nucleico que comprende un gen sintético que permite una expresión optimizada de la proteína S, vector que es conocido en la cepa bacteriana depositada en la CNCM, el 1 de diciembre de 2004, bajo el n° I-3333.

Según un modo de realización de dicho vector de expresión, se trata de un vector viral, en forma de partícula viral o en forma de genoma recombinante.

Según una disposición ventajosa de este modo de realización, se trata de una partícula viral recombinante o de un genoma viral recombinante susceptible de ser obtenido por transfección de un plásmido según los apartados g), h) y k) a r), tal como se han definido anteriormente, en un sistema celular apropiado, es decir, por ejemplo, células transfectadas con uno o algunos otros plásmido(s), destinados a transcomplementar ciertas funciones del virus eliminadas en el vector y necesarias para la formación de las partículas virales.

Se entiende aquí por "familia de la proteína S" la proteína S completa, su ectodominio y fragmentos de este ectodominio, que se producen preferiblemente en un sistema eucariota.

Los inventores describen también un vector lentiviral que codifica un polipéptido de la familia de la proteína S, tal como se ha definido anteriormente.

Los inventores describen también un virus de la rubeola recombinante, que codifica un polipéptido de la familia de la proteína S, tal como se ha definido anteriormente.

Los inventores describen también un virus de la vacuna recombinante que codifica un polipéptido de la familia de la proteína S, tal como se ha definido anteriormente.

Los inventores describen también el uso de un vector según los apartados e) a r) tales como se han definido anteriormente, o de un vector que comprende un gen sintético de la proteína S, tal como se ha definido anteriormente, para la producción, en un sistema eucariota, de la proteína S del coronavirus asociado al SRAS, o de un fragmento de esta proteína.

Los inventores describen también un método de producción de la proteína S en un sistema eucariota, que comprende una etapa de transfección de células eucariotas en cultivo por un vector seleccionado entre los vectores contenidos en las cepas bacterianas mencionadas en los apartados e) a r) anteriores, o un vector que comprende un gen sintético que permite una expresión optimizada de la proteína S.

Los inventores describen también un banco de ADNc caracterizado porque comprende unos fragmentos tales como se han definido anteriormente, en particular unos fragmentos de amplificación o unos fragmentos de restricción, clonados en un vector recombinante, en particular un vector de expresión (banco de expresión).

Los inventores describen también unas células, en particular unas células procariotas, modificadas por un vector recombinante tal como se ha definido anteriormente.

Los inventores describen también una célula eucariota genéticamente modificada, que expresa una proteína o un polipéptido tales como se definen más adelante. Por supuesto, los términos "células eucariota genéticamente modificada" no designan una célula modificada por un virus salvaje.

Según un modo de realización ventajoso de dicha célula, es susceptible de ser obtenida mediante transfección por cualquiera de los vectores mencionados en los apartados k) a n) anteriores.

Según una disposición ventajosa de este modo de realización, se trata de la célula FRhK4-Ssol-30, depositada en la CNCM el 22 de noviembre de 2004, bajo el nº I-3325.

Los vectores recombinantes tales como se han definido anteriormente y las células transformadas por dichos vectores de expresión son ventajosamente utilizados para la producción de proteínas y de péptidos correspondientes. Los bancos de expresión derivados de dichos vectores, así como las células transformadas por dichos bancos de expresión, son ventajosamente utilizados para identificar los epítomos inmunógenos (epítomos B y T) de las proteínas del coronavirus asociado al SRAS.

Los inventores describen también las proteínas y los péptidos purificados o aislados, caracterizados porque están codificados por el polinucleótido o uno de sus fragmentos tales como se han definido anteriormente.

Según un modo de realización ventajoso, dicha proteína se selecciona del grupo constituido por:

- la proteína S de secuencia SEC ID nº 3 o su ectodominio
- la proteína E de secuencia SEC ID nº 14
- la proteína M de secuencia SEC ID nº 17
- la proteína N de secuencia SEC ID nº 37
- las proteínas codificadas por ORF: ORF1a, ORF1b, ORF3, ORF4 y ORF7 a ORF11, ORF 13 y ORF 14 de secuencia respectivamente, SEC ID nº 74, 75, 10, 12, 22, 24, 26, 28, 30, 33 y 35.

Se usarán a continuación, indiferentemente, los términos "ectodominio de la proteína S" y "forma soluble de la proteína S".

Según un modo de realización ventajoso, dicho polipéptido está constituido de aminoácidos que corresponden a las posiciones 1 a 1193 de la secuencia en aminoácidos de la proteína S.

Según otro modo de realización ventajoso, dicho péptido se selecciona del grupo constituido por:

- a) los péptidos que corresponden a las posiciones 14 a 1193 y 475 a 1193 de la secuencia en aminoácidos de la proteína S,
- b) los péptidos que corresponden a las posiciones 2 a 14 (SEC ID nº 69) y 100 a 221 de la secuencia en aminoácidos de la proteína M; estos péptidos corresponden respectivamente al ectodominio y al endodominio de la proteína M, y
- c) los péptidos que corresponden a las posiciones 1 a 12 (SEC ID nº 70) y 53 a 76 (SEC ID nº 71) de la secuencia en aminoácidos de la proteína E; estos péptidos corresponden respectivamente al ectodominio y al extremo C-terminal de la proteína E, y
- d) los péptidos de 5 a 50 aminoácidos consecutivos, preferiblemente de 10 a 30 aminoácidos, incluidos o que se superponen parcial o totalmente a la secuencia de los péptidos tales como se han definido en a), b) o c).

Los inventores describen también un péptido caracterizado porque presenta una secuencia de 7 a 50 aminoácidos, que incluye un residuo de aminoácido seleccionado del grupo constituido por:

- la alanina, situada en la posición 2552 de la secuencia en aminoácidos de la proteína codificada por el ORF1a,
- la serina, situada en la posición 577 de la secuencia en aminoácidos de la proteína S de la cepa de SRAS-CoV tal como se ha definido anteriormente,
- la glicina, en la posición 11 de la secuencia en aminoácidos de la proteína codificada por el ORF3 de la cepa de SRAS-CoV tal como se ha definido anteriormente,
- la serina, en la posición 154 de la secuencia en aminoácidos de la proteína M de la cepa de SRAS-CoV tal como se ha definido anteriormente,

Los inventores describen también un anticuerpo o un fragmento de anticuerpo policlonal o monoclonal, susceptible

de ser obtenido mediante inmunización de un animal con un vector recombinante, tal como se ha definido anteriormente, un banco de ADNc tal como se ha definido anteriormente o bien una proteína o un péptido tales como se han definido anteriormente, caracterizado porque se enlaza con al menos una de las proteínas codificadas por el SRAS-CoV tal como se ha definido anteriormente.

Un anticuerpo, tal como se ha definido anteriormente, engloba los anticuerpos policlonales, los anticuerpos monoclonales, los anticuerpos quiméricos, tales como los anticuerpos humanizados, así como sus fragmentos (Fab, Fv, scFv).

Los inventores describen también un hibridoma, que produce un anticuerpo monoclonal contra la proteína N, caracterizado porque se selecciona entre los hibridomas siguientes:

- el hibridoma que produce el anticuerpo monoclonal 87, depositado en la CNCM el 1 de diciembre de 2004 bajo el número I-3328,
- el hibridoma que produce el anticuerpo monoclonal 86, depositado en la CNCM el 1 de diciembre de 2004 bajo el número I-3329,
- el hibridoma que produce el anticuerpo monoclonal 57, depositado en la CNCM el 1 de diciembre de 2004 bajo el número I-3330, y
- el hibridoma que produce el anticuerpo monoclonal 156, depositado en la CNCM el 1 de diciembre de 2004 bajo el número I-3331.

Los inventores describen también un anticuerpo, o fragmento de anticuerpo policlonal o monoclonal, dirigido contra la proteína N, caracterizado porque está producido por un hibridoma tal como se ha definido anteriormente.

Se entiende por anticuerpo quimérico, en relación a un anticuerpo de una especie animal particular o de una clase particular de anticuerpo, un anticuerpo que comprende todo o parte de una cadena pesada y/o de una cadena ligera de un anticuerpo de otra especie animal o de otra clase de anticuerpo.

Se entiende por anticuerpo humanizado una inmunoglobulina humana en la que los residuos de CDR (*Complementary-Determining Regions*), que forman el sitio de unión al antígeno, están sustituidos por los de un anticuerpo monoclonal no humano que posee la especificidad, la afinidad o la actividad buscadas. Por comparación con los anticuerpos no humanos, los anticuerpos humanizados son menos inmunógenos y poseen una vida media prolongada en el ser humano, ya que posee sólo una baja proporción de secuencias no humanas, debido a que la casi totalidad de los residuos de las regiones FR (Framework) y de la región constante (Fc) de estos anticuerpos son los de unas secuencias consenso de inmunoglobulinas humanas.

Los inventores describen también un chip o filtro de proteína, caracterizado porque comprende una proteína, un péptido o bien un anticuerpo tales como se han definido anteriormente.

Los chips de proteína son preparados mediante los métodos clásicos, conocidos. Entre los soportes apropiados sobre los cuales se pueden inmovilizar unas proteínas, se pueden citar los de material plástico o de vidrio, en particular en forma de microplacas.

Los inventores describen también agentes reactivos derivados de la cepa aislada de coronavirus asociado al SRAS, procedente de la extracción catalogada bajo el n° 031589, útiles para el estudio y el diagnóstico de la infección provocada por un coronavirus asociado al SRAS, agentes reactivos que se seleccionan del grupo constituido por:

- (a) un par de cebadores, una sonda o un chip de ADN tales como se han definido anteriormente,
- (b) un vector recombinante, o una célula modificada, tales como se han definido anteriormente,
- (c) una cepa aislada de coronavirus, o un polinucleótido, tales como se han definido anteriormente,
- (d) una proteína o un péptido tales como se han definido anteriormente,
- (e) un anticuerpo o un fragmento de anticuerpo, tales como se han definido anteriormente,
- (f) un chip de proteína tal como se ha definido anteriormente.

Estos diferentes agentes reactivos son preparados y utilizados según las técnicas clásicas de biología molecular y de inmunología, siguiendo los protocolos estándares, tales como los descritos en *Current Protocols in Molecular Biology* (Frederick M. AUSUBEL, 2000, Wiley and Son Inc., Library of Congress, USA), en *Current Protocols in Immunology* (John E. Cologan, 2000, Wiley and Son Inc. Library of Congress, USA) y en *Antibodies: A Laboratory*

Manual (E. Howell y D Lane, Cold Spring Harbor Laboratory, 1988).

Los fragmentos de ácido nucleico, según la invención, son preparados y utilizados según las técnicas clásicas, tales como se han definido anteriormente. Los péptidos y las proteínas, según la invención, son preparados mediante las técnicas de ADN recombinante, conocidas por el experto en la materia, en particular con la ayuda de los vectores recombinantes tales como se han definido anteriormente. Alternativamente, los péptidos, según la invención, se pueden preparar mediante las técnicas clásicas de síntesis en fase sólida o líquida, conocidas por el experto en la materia.

Los anticuerpos policlonales son preparados mediante inmunización de un animal apropiado con una proteína o un péptido tales como se han definido anteriormente, eventualmente acoplado a la KLH o a la albúmina y/o asociado a un adyuvante apropiado tal como el adyuvante de Freund (completo o incompleto) o el hidróxido de alúmina; después de la obtención de un título en anticuerpo satisfactorio, los anticuerpos se recogen mediante extracción del suero de los animales inmunizados y enriquecidos en IgG por precipitación, según las técnicas clásicas, después de los IgG específicos de las proteínas de SRAS-CoV se purifican eventualmente mediante cromatografía de afinidad sobre una columna apropiada, sobre la cual se fijan dicho péptido o dicha proteína, tales como se han definido anteriormente, a fin de obtener una preparación de IgG monoespecíficos.

Los anticuerpos monoclonales son producidos a partir de hibridomas obtenidos mediante fusión de linfocitos B de un animal inmunizado por una proteína o un péptido tales como se han definido anteriormente con unos mielomas, según la técnica de Kohler y Milstein (Nature, 1975, 256, 495-497); los hibridomas son cultivados *in vitro*, en particular en fermentadores o producidos *in vivo*, en forma de ascito; alternativamente, dichos anticuerpos monoclonales son producidos mediante ingeniería genética tal, como se describe en la patente americana US 4.816.567.

Los anticuerpos humanizados son producidos mediante métodos generales como los descritos en la solicitud internacional WO 98/45332.

Los fragmentos de anticuerpos son producidos a partir de las regiones V_H y V_L clonadas, a partir de los ARNm de hibridomas o de linfocitos esplénicos de un ratón inmunizado; por ejemplo, los fragmentos Fv, scFv o Fab están expresados en la superficie de fagos filamentosos según la técnica de Winter y Milstein (Nature, 1991, 349, 293-299); después de varias etapas de selección, los fragmentos de anticuerpos específicos del antígeno son aislados y expresados en un sistema de expresión apropiado, mediante las técnicas clásicas de clonación y de expresión de ADN recombinante.

Los anticuerpos o sus fragmentos tales como se han definido anteriormente, son purificados mediante las técnicas clásicas conocidas por el experto en la materia, tales como la cromatografía de afinidad.

Los inventores describen además el uso de un producto seleccionado del grupo constituido por: un par de cebadores, una sonda, un chip de ADN, un vector recombinante, una célula modificada, una cepa aislada de coronavirus, un polinucleótido, una proteína o un péptido, un anticuerpo o un fragmento de anticuerpo y un chip de proteína tales como se han definido anteriormente, para la preparación de un agente reactivo de detección y eventualmente de genotipaje/serotipaje, de un coronavirus asociado al SRAS.

Las proteínas y los péptidos tales como se han definido anteriormente, que son adecuados para ser reconocidos y/o para inducir la producción de anticuerpos específicos del coronavirus asociado al SRAS, son útiles para el diagnóstico de la infección por tal coronavirus; la infección se detecta, mediante una técnica apropiada, en particular EIA, ELISA, RIA, inmunofluorescencia, a partir de una muestra biológica extraída en un individuo susceptible de ser infectado.

Según una disposición ventajosa de dicho uso, dichas proteínas se seleccionan del grupo constituido por las proteínas S, E, M y/o N, y los péptidos tales como se han definido anteriormente.

Las proteínas S, E, M y/o N y los péptidos derivados de estas proteínas tales como se han definido anteriormente, por ejemplo la proteína N, se utilizan para el diagnóstico indirecto de una infección con coronavirus asociado al SRAS (diagnóstico serológico; detección de anticuerpos específicos del SRAS-CoV), en particular mediante un método inmunoenzimático (ELISA).

Los anticuerpos y los fragmentos de anticuerpos tales como se han definido anteriormente, en particular los dirigidos contra las proteínas S, E, M y/o N y los péptidos derivados tales como se han definido anteriormente, son útiles para el diagnóstico directo de una infección con coronavirus asociado al SRAS; la detección de proteína(s) del SRAS-CoV se realiza mediante una técnica apropiada, en particular EIA, ELISA, RIA, inmunofluorescencia a partir de una muestra biológica extraída en un individuo susceptible de ser infectado.

Los inventores describen asimismo un método de detección de un coronavirus asociado al SRAS, a partir de una muestra biológica, método que se caracteriza porque comprende al menos:

- (a) la puesta en contacto de dicha muestra biológica con al menos un anticuerpo o un fragmento de anticuerpo, una proteína, un péptido o bien un chip o un filtro de proteína o de péptido, tales como se han definido anteriormente, y
- (b) el revelado mediante cualquier medio apropiado unos complejos antígeno-anticuerpo formados en (a), por ejemplo mediante EIA, ELISA, RIA, o mediante inmunofluorescencia.

Según un modo de aplicación ventajoso de dicho procedimiento, la etapa (a) comprende:

- (a₁) la puesta en contacto de dicha muestra biológica con al menos un primer anticuerpo o un fragmento de anticuerpo que está fijado sobre un soporte apropiado, en particular una microplaca,
- (a₂) el lavado de la fase sólida, y
- (a₃) la adición de al menos un segundo anticuerpo o un fragmento de anticuerpo, diferente del primero, estando dicho anticuerpo o fragmento de anticuerpo eventualmente marcado de manera apropiada.

Este procedimiento que permite capturar las partículas víricas presentes en la muestra biológica se denomina también procedimiento de inmunocaptura.

Por ejemplo:

- la etapa (a₁) se realiza con al menos un primer anticuerpo monoclonal o policlonal o un fragmento de estos, dirigido contra la proteína S, M y/o E, y/o un péptido que corresponde al ectodominio de una de estas proteínas (péptidos M2-14 o E1-12)

- la etapa (a₃) se realiza con al menos un anticuerpo o un fragmento de anticuerpo dirigido contra otro epítipo de la misma proteína o preferiblemente contra otra proteína, de manera preferente contra una proteína interna, tal como la nucleoproteína N o el endodominio de la proteína E o M, de manera aún más preferente, se trata de anticuerpos o de fragmentos de anticuerpos dirigidos contra la proteína N, muy abundante en la partícula viral; cuando un anticuerpo o un fragmento de anticuerpo dirigido contra una proteína interna (N) o contra el endodominio de las proteínas E o M se utiliza, dicho anticuerpo es incubado en presencia de detergente, como el Tween 20 por ejemplo, a unas concentraciones del orden del 0,1%,

- la etapa (b) de revelado de los complejos antígeno-anticuerpo formados se realiza directamente con la ayuda de un segundo anticuerpo marcado, por ejemplo, con biotina o una enzima apropiada tal como la peroxidasa o la fosfatasa alcalina, o bien indirectamente con la ayuda de un suero anti-inmunoglobulinas marcado como anteriormente. Los complejos así formados son revelados con la ayuda de un sustrato apropiado.

Según una realización preferente de este método, la muestra biológica se mezcla con el anticuerpo monoclonal de revelado previamente a su puesta en contacto con los anticuerpos monoclonales de captura. Llegado el caso, la mezcla suero-anticuerpo de revelado se incuba durante al menos 10 minutos a temperatura ambiente antes de ser aplicada sobre la placa.

Los inventores describen también un ensayo de inmunocaptura destinado a detectar una infección por el coronavirus asociado al SRAS por detección de la nucleoproteína nativa (proteína N), en particular caracterizado porque el anticuerpo utilizado para la captura de la nucleoproteína viral nativa es un anticuerpo monoclonal específico de la región central y/o de un epítipo conformacional.

Según un modo de realización de dicho ensayo, el anticuerpo utilizado para la captura de la proteína N es el anticuerpo monoclonal Acm87, producido por el hibridoma depositado en la CNCM el 1 de diciembre de 2004 bajo el número I-3328.

Según otro modo de realización de dicho ensayo de inmunocaptura, el anticuerpo utilizado para la captura de la proteína N es el anticuerpo monoclonal Acm86, producido por el hibridoma depositado en la CNCM el 1 de diciembre de 2004 bajo el número I-3329.

Según otro modo de realización de dicho ensayo de inmunocaptura, los anticuerpos monoclonales Acm86 y Acm87 son utilizados para la captura de la proteína N.

En los ensayos de inmunocaptura, se puede utilizar el revelado de la proteína N, el anticuerpo monoclonal Acm57, producido por el hibridoma depositado en la CNCM el 1 de diciembre de 2004 bajo el número 1-3330, estando dicho anticuerpo conjugado a una molécula o a una partícula reveladora.

Conforme a dicho ensayo de inmunocaptura, una combinación de los anticuerpos Acm57 y Acm87, conjugados a

una molécula o a una partícula reveladora, se utiliza para el revelado de la proteína N.

Una molécula reveladora puede ser un átomo radioactivo, un colorante, una molécula fluorescente, un fluoróforo, una enzima; una partícula reveladora puede ser, por ejemplo, oro coloidal, una partícula magnética o una bola de látex.

Los inventores describen también un agente reactivo de detección de un coronavirus asociado al SRAS, caracterizado porque se selecciona del grupo constituido por:

- (a) un par de cebadores o una sonda tales como se han definido anteriormente,
- (b) un vector recombinante tal como se ha definido anteriormente o una célula modificada tal como se ha definido anteriormente,
- (c) una cepa aislada de coronavirus tal como se ha definido anteriormente o un polinucleótido tal como se ha definido anteriormente,
- (d) un anticuerpo o un fragmento de anticuerpo tal como se ha definido anteriormente,
- (e) una combinación de anticuerpos que comprende los anticuerpos monoclonales Acm86 y/o Acm87, y el anticuerpo monoclonal Acm57, tal como se ha definido anteriormente,
- (f) un chip o un filtro, tal como se han definido anteriormente.

Los inventores describen también un método de detección de una infección por un coronavirus asociado al SRAS, a partir de una muestra biológica, mediante ELISA IgG indirecto que utiliza la proteína N, método que está caracterizado porque las placas son sensibilizadas por una solución de proteína N a una concentración comprendida entre 0,5 y 4 µg/ml, preferiblemente 2 µg/ml, en un tampón PBS 10 mM pH 7,2, rojo de fenol a 0,25 ml/l.

Los inventores describen, además, un método de detección de una infección por un coronavirus asociado al SRAS, a partir de una muestra biológica, mediante ELISA doble epítipo, caracterizado porque el suero a ensayar está mezclado al antígeno de revelado, siendo dicha mezcla puesta en contacto después con el antígeno fijado sobre un soporte sólido.

Según una variante de los ensayos de detección de los coronavirus asociado al SRAS, estos ensayos combinan un ELISA que utiliza la proteína N, y otro ELISA que utiliza la proteína S, tal como se describe más adelante.

Los inventores describen también un complejo inmune formado de un anticuerpo o de un fragmento de anticuerpo policlonal o monoclonal, tal como se han definido anteriormente, y de una proteína o de un péptido del coronavirus asociado al SRAS.

Los inventores describen además un kit de detección de un coronavirus asociado al SRAS, caracterizado porque comprende al menos un agente reactivo seleccionado del grupo constituido por: un par de cebadores, una sonda, un chip de ADN o ARN, un vector recombinante, una célula modificada, una cepa aislada de coronavirus, un polinucleótido, una proteína o un péptido, un anticuerpo, y un chip de proteína, tales como se han definido anteriormente.

Los inventores describen además una composición inmunógena, caracterizada porque comprende al menos un producto seleccionado del grupo constituido por:

- a) una proteína o un péptido tales como se han definido anteriormente,
- b) un polinucleótido de tipo ADN o ARN o uno de sus fragmentos representativos tales como se han definido anteriormente, de secuencia seleccionada entre:
 - (i) la secuencia SEC ID nº 1 o su equivalente ARN,
 - (ii) la secuencia que hibrida, en condiciones de fuerte rigor, con la secuencia SEC ID nº 1,
 - (iii) la secuencia complementaria de la secuencia SEC ID nº 1 o de la secuencia que hibrida en condiciones de fuerte rigor con la secuencia SEC ID nº 1,
 - (iv) la secuencia nucleotídica de un fragmento representativo del polinucleótido, tal como se define en (i), (ii) o (iii),
 - (v) la secuencia tal como se define en (i), (ii), (iii) o (iv), modificada, y

- c) un vector de expresión recombinante que comprende un polinucleótido tal como se define en b), y
- d) un banco de ADNc tal como se definió anteriormente,

siendo dicha composición inmunógena capaz de inducir una inmunidad humoral o celular protectora específica del coronavirus asociado al SRAS, en particular la producción de un anticuerpo dirigido contra un epítipo específico del coronavirus asociado al SRAS.

Las proteínas y los péptidos, tales como se han definido anteriormente, en particular las proteínas S, M, E y/o N y los péptidos derivados, así como las moléculas de ácido nucleico (ADN o ARN) que codifican dichas proteínas o dichos péptidos, son buenas vacunas candidatas y pueden ser utilizadas en composiciones inmunógenas para la producción de una vacuna contra el coronavirus asociado al SRAS.

Según un modo de realización ventajoso de las composiciones, según la invención, estas contienen además al menos un vehículo farmacéuticamente aceptable y eventualmente sustancias portadoras y/o adyuvantes.

Los vehículos farmacéuticamente aceptables, las sustancias portadoras y los adyuvantes son los clásicamente utilizados.

Los adyuvantes se seleccionan ventajosamente del grupo constituido por emulsiones oleosas, saponina, sustancias minerales, extractos bacterianos, hidróxido de alúmina y escualeno.

Las sustancias portadoras son ventajosamente seleccionadas del grupo constituido por los liposomas unilaminares, los liposomas multilaminares, unas micelas de saponina o unas microesferas sólidas de naturaleza sacarídica o aurífera.

Las composiciones, según la invención, se administran por vía general, en particular intramuscular o subcutánea, o bien por vía local, en particular nasal (aerosol).

Los inventores describen también el uso de una proteína o de un péptido aislado o purificado que presenta una secuencia seleccionada del grupo constituido por las secuencias SEC ID nº 3, 10, 12, 14, 17, 22, 24, 26, 28, 30, 33, 35, 37, 69, 70, 71, 74 y 75 para formar un complejo inmune con un anticuerpo dirigido específicamente contra un epítipo del coronavirus asociado al SRAS.

Los inventores describen también un complejo inmune formado por una proteína o por un péptido aislado o purificado, que presenta una secuencia seleccionada del grupo constituido por las secuencias SEC ID nº 3, 10, 12, 14, 17, 22, 24, 26, 28, 30, 33, 35, 37, 69, 70, 71, 74 y 75, y de un anticuerpo dirigido específicamente contra un epítipo del coronavirus asociado al SRAS.

Los inventores describen también el uso de una proteína o de un péptido aislado o purificado, que presenta una secuencia seleccionada del grupo constituido por las secuencias SEC ID nº 3, 10, 12, 14, 17, 22, 24, 26, 28, 30, 33, 35, 37, 69, 70, 71, 74 y 75, para inducir la producción de un anticuerpo capaz de reconocer específicamente un epítipo del coronavirus asociado al SRAS.

Los inventores describen también el uso de un polinucleótido aislado o purificado, que presenta una secuencia seleccionada del grupo constituido por las secuencias SEC ID nº 1, 2, 4, 7, 8, 13, 15, 16, 18, 19, 20, 31, 36 y 38 para inducir la producción de un anticuerpo dirigido contra la proteína codificada por dicho polinucleótido, y capaz de reconocer específicamente un epítipo del coronavirus asociado al SRAS.

Los inventores describen también unos anticuerpos monoclonales que reconocen la proteína S nativa de un coronavirus asociado al SRAS.

Los inventores describen también el uso de una proteína o de un polipéptido de la familia de la proteína S, tal como se ha definido anteriormente, o de un anticuerpo que reconoce la proteína S nativa, tal como se ha definido anteriormente, para detectar una infección por un coronavirus asociado al SRAS, a partir de una muestra biológica.

Los inventores describen también un método de detección de una infección por un coronavirus asociado al SRAS, a partir de una muestra biológica, caracterizado porque la detección se efectúa mediante ELISA utilizando la proteína S recombinante, expresada en un sistema eucariota.

Según un modo de aplicación ventajoso de dicho método, se trata de un método por ELISA doble epítipo, y el suero a ensayar se mezcla con el antígeno de revelado, siendo dicha mezcla puesta en contacto después con el antígeno fijado sobre un soporte sólido.

Los inventores describen también un complejo inmune formado de un anticuerpo o de un fragmento de anticuerpo

monoclonal que reconoce la proteína S nativa, y de una proteína o de un péptido del coronavirus asociado al SRAS.

Los inventores describen también un complejo inmune formado de una proteína o de un polipéptido de la familia de la proteína S, tal como se ha definido anteriormente, y de un anticuerpo dirigido específicamente contra un epítipo del coronavirus asociado al SRAS.

Los inventores describen además un kit o estuche de detección de un coronavirus asociado al SRAS, caracterizado porque comprende al menos un agente reactivo seleccionado del grupo constituido por: una proteína o polipéptido de la familia de la proteína S, tal como se ha definido anteriormente, un ácido nucleico que codifica una proteína o un polipéptido de la familia de la proteína S, tal como se ha definido anteriormente, una célula que expresa una proteína o un polipéptido de la familia de la proteína S, tal como se ha definido anteriormente, o un anticuerpo que reconoce la proteína S nativa de un coronavirus asociado al SRAS.

Los inventores describen una composición inmunógena y/o de vacuna, caracterizada porque comprende un polipéptido o una proteína recombinante de la familia de la proteína S, tal como se ha definido anteriormente, obtenido en un sistema de expresión eucariota.

Los inventores describen también una composición inmunógena y/o de vacuna, caracterizada porque comprende un vector o virus recombinante, que expresa una proteína o un polipéptido de la familia de la proteína S, tal como se ha definido anteriormente.

Además de las disposiciones anteriores, la invención comprende también otras disposiciones, que destacarán de la descripción siguiente, que se refiere a unos ejemplos de aplicación del polinucleótido que representa el genoma de la cepa de SRAS-CoV procedente de la extracción catalogada bajo el número 031589, y unos fragmentos de ADNc derivados, que son objeto de la presente invención, así como a la Tabla I que presenta el listado de las secuencias:

Tabla I: listado de secuencias

Número de identificación	Secuencia	Posición del ADNc en referencia a Genbank AY274119.3	Número de depósito en la CNCM del plásmido correspondiente
SEC ID nº 1	Genoma de la cepa procedente de la extracción 031589	-	-
SEC ID nº 2	ORF-S*	21406-25348	-
SEC ID nº 3	Proteína S	-	-
SEC ID nº 4	ORF-S**	21406-25348	I-3059
SEQ ID nº 5	fragmento Sa	21406-23454	I-3020
SEC ID nº 6	fragmento Sb	23322-25348	I-3019
SEC ID nº 7	ORF-3+ORF-4*	25110-26244	-
SEC ID nº 8	ORF-3+ORF-4**	25110-26244	I-3126
SEC ID nº 9	ORF3	-	-
SEC ID nº 10	Proteína ORF-3	-	-
SEC ID nº 11	ORF4	-	-
SEC ID nº 12	Proteína ORF-4	-	-
SEQ ID nº 13	ORF-E*	26082-26413	-
SEC ID nº 14	Proteína E	-	-
SEC ID nº 15	ORF-E**	26082-26413	I-3046
SEC ID nº 16	ORF-M*	26330-27098	-
SEC ID nº 17	Proteína M	-	-
SEC ID nº 18	ORF-M**	26330-27098	I-3047
SEC ID nº 19	ORF7 a 11*	26977-28218	-
SEC ID nº 20	ORF7 a 11**	26977-28218	I-3125
SEC ID nº 21	ORF7	-	-
SEC ID nº 22	Proteína ORF7	-	-
SEC ID nº 23	ORF8	-	-
SEC ID nº 24	Proteína ORF8	-	-
SEC ID nº 25	ORF9	-	-
SEC ID nº 26	Proteína ORF9	-	-
SEC ID nº 27	ORF10	-	-
SEC ID nº 28	Proteína ORF10	-	-
SEC ID nº 29	ORF11	-	-
SEC ID nº 30	Proteína ORF11	-	-

Número de identificación	Secuencia	Posición del ADNc en referencia a Genbank AY274119.3	Número de depósito en la CNCM del plásmido correspondiente
SEC ID nº 31	OrF1ab	265-21485	-
SEC ID nº 32	ORF13	28130-28426	-
SEC ID nº 33	Proteína ORF13	-	-
SEC ID nº 34	ORF14	-	-
SEC ID nº 35	Proteína ORF14	28583-28795	-
SEC ID nº 36	ORF-N*	28054-29430	-
SEC ID nº 37	Proteína N	-	-
SEC ID nº 38	ORF-N**	28054-29430	I-3048
SEC ID nº 39	5'no codificante**	1-204	I-3124
SEC ID nº 40	3'no codificante**	28933-29727	I-3123
SEC ID nº 41	ORF1ab Fragmento L0	30-500	-
SEC ID nº 42	Fragmento L1	211-2260	-
SEC ID nº 43	Fragmento L2	2136-4187	-
SEC ID nº 44	Fragmento L3	3892-5344	-
SEC ID nº 45	Fragmento L4b	4932-6043	-
SEC ID nº 46	Fragmento L4	5305-7318	-
SEC ID nº 47	Fragmento L5	7275-9176	-
SEC ID nº 48	Fragmento L6	9032-11086	-
SEC ID nº 49	Fragmento L7	10298-12982	-
SEC ID nº 50	Fragmento L8	12815-14854	-
SEC ID nº 51	Fragmento L9	14745-16646	-
SEC ID nº 52	Fragmento L10	16514-18590	-
SEC ID nº 53	Fragmento L11	18500-20602	-
SEC ID nº 54	Fragmento L12	20319-22224	-
SEC ID nº 55	Cebador N sentido	-	-
SEC ID nº 56	Cebador N antisentido	-	-
SEC ID nº 57	Cebador S _C sentido	-	-
SEC ID nº 58	Cebador S _L sentido	-	-
SEC ID nº 59	Cebador S _C y S _L antisentido	-	-
SEC ID nº 60	Cebador sentido serie 1	28507-28522	-
SEC ID nº 61	Cebador antisentido serie 1	28774-28759	-
SEC ID nº 62	Cebador sentido serie 2	28375-28390	-
SEC ID nº 63	Cebador antisentido serie 2	28702-28687	-
SEC ID nº 64	Sonda 1/serie 1	28561-28586	-
SEC ID nº 65	Sonda 2/serie 1	28588-28608	-
SEC ID nº 66	Sonda 1/serie 2	28541-28563	-
SEC ID nº 67	Sonda 2/serie 2	28565-28589	-
SEC ID nº 68	Cebador ancla 14T		
SEC ID nº 69	Péptido M2-14	-	-
SEC ID nº 70	Péptido E1-12	-	-
SEC ID nº 71	Péptido E53-76	-	-
SEC ID nº 72	5' no codificante*	1-204	-
SEC ID nº 73	3' no codificante*	28933-29727	-
SEC ID nº 74	Proteína ORF1a	-	-
SEC ID nº 75	Proteína ORF1b	-	-
SEQ ID nº ^{os} 76-139	Cebadores		
SEQ ID nº 140	Pseudógeno de S		
SEQ ID nº ^{os} 141-148	cebadores		
SEQ ID nº 149	Aa1-13 de S		
SEQ ID nº 150	polipéptido		
SEQ ID nº ^{os} 151-158	cebadores		

* producto de amplificación PCR (amplicón)
 ** inserto clonado en el plásmido depositado en la CNCM

así como a los dibujos anexos en los que:

- la figura 1 ilustra el análisis por transferencia western de la expresión *in vitro* de las proteínas recombinantes N,

S_C y S_L a partir de los vectores de expresión pIVEX. Pista 1: pIV2.3N. Pista 2: pIV2.3S_C. Pista 3: pIV2.3S_L. Pista 4: pIV2.4N. Pista 5: pIV2.4S₁ o pIV2.4S_C. Pista 6: pIV2.4S_L. La expresión de la proteína GFP expresada a partir del mismo vector se utiliza como control.

- la figura 2 ilustra el análisis por electroforesis en gel de poliacrilamida en condiciones desnaturalizantes (SDS-PAGE), y coloración con azul de Coomassie, de la expresión *in vivo* de la proteína N a partir de los vectores de expresión pIVEX. La cepa de *E.coli* BL21(DE3)pDIA17, transformada por los vectores pIVEX recombinantes, se cultiva a 30°C en un medio LB, en presencia o ausencia de inductor (IPTG 1mM). Pista 1: pIV2.3N Pista 2: pIV2.4N.
- la figura 3 ilustra el análisis por electroforesis en gel de poliacrilamida en condiciones desnaturalizantes (SDS-PAGE) y coloración con azul de Coomassie, de la expresión *in vivo* de los polipéptidos S_L y S_C a partir de los vectores de expresión pIVEX. La cepa de *E.coli* BL21 (DE3)pDIA17, transformada por los vectores pIVEX recombinantes, se cultiva a 30°C en un medio LB, en presencia o ausencia de inductor (IPTG 1mM). Pista 1: pIV2.3S_C Pista 2: pIV2.3S_L. Pista 3: pIV2.4S₁ Pista 4: pIV2.4S_L.
- la figura 4 ilustra la actividad antigénica de las proteínas N, S_L y S_C recombinantes producidas en la cepa *E. coli* BL21(DE3)pDIA17, transformada por los vectores pIVEX recombinantes. A: electroforesis (SDS-PAGE) de los lisados bacterianos. B y C: transferencia western con los sueros, que provienen de un mismo paciente infectado por SRAS-CoV, extraídos respectivamente a los 8 días (B: suero M12) y a los 29 días (C: suero M13) después del comienzo de los síntomas del SRAS. Pista 1: pIV2.3N. Pista 2: pIV2.4N. Pista 3: pIV2.3S_C. Pista 4: pIV2.4S₁. Pista 5: pIV2.3S_L. Pista 6: pIV2.4S_L.
- la figura 5 ilustra la purificación sobre columna Ni-NTA agarosa de la proteína N recombinante, producida en la cepa *E. coli* BL21(DE3)pDIA17 a partir del vector pIV2.3N. Pista 1: Extracto bacteriano total. Pista 2: Extracto soluble. Pista 3: Extracto insoluble. Pista 4: Extracto depositado sobre la columna Ni-NTA. Pista 5: proteínas no retenidas. Pista 6: Fracciones del pico 1. Pista 7: Fracciones del pico 2.
- la figura 6 ilustra la purificación de la proteína S_C recombinante a partir de los cuerpos de inclusiones producidos en la cepa *E. coli* BL21(DE3)pDIA17, transformada por el pIV2.4S₁. A. Tratamiento con Tritón X-100 (2%): Pista 1: Extracto bacteriano total. Pista 2: Extracto soluble. Pista 3: Extracto insoluble. Pista 4: Sobrenadante después del tratamiento con Tritón X-100 (2%). Pistas 5 y 6: Residuo después del tratamiento con Tritón X-100 (2%). B: Tratamiento con urea 4M, 5M, 6M y 7M de los extractos solubles e insolubles.
- la figura 7 representa la inmunohuella realizada con la ayuda de un lisado de células infectadas por el SRAS-CoV y de un suero de paciente que padece neumopatía atípica.
- la figura 8 representa unas inmunohuellas realizadas con la ayuda de un lisado de células infectadas por el SRAS-CoV y de inmunosueros de conejos específicos de la nucleoproteína N (A) y de la proteína de espícula S (B). I.S.: suero inmune. p.i.: suero pre-inmune. El inmunosero anti-N se utilizó al 1/50000 y el inmunosero anti-S al 1/10000.
- la figura 9 ilustra la reactividad en ELISA de los sueros policlonales mono-específicos de conejo dirigidos contra la proteína N o el fragmento corto de la proteína S (S_C), frente a proteínas recombinantes correspondientes utilizadas para la inmunización. A: conejos P13097, P13081, y P13031 inmunizados con la proteína N recombinante purificada. B: conejos P11135, P13042, y P14001 inmunizados con una preparación de cuerpos de inclusiones que corresponden al fragmento corto de la proteína S (S_C). I.S.: suero inmune. p.i.: suero pre-inmune.
- la figura 10 ilustra la reactividad en ELISA de la proteína N recombinante purificada, frente a suero de pacientes que padecen neumonía atípica causada por el SRAS-CoV. Figura 10a: placas ELISA preparadas con la proteína N a la concentración de 4 µg/ml y 2 µg/ml. Figura 10b: placa ELISA preparada con la proteína N a la concentración de 1 µg/ml. Los sueros designados A, B, D, E, F, G, H corresponden a los de la Tabla IV.
- la figura 11 ilustra la amplificación por RT-PCR de cantidades decrecientes de ARN sintético del gen N del SRAS-CoV (10⁷ a 1 copia), con la ayuda de los pares de cebadores n° 1 (N/+28507,N/-28774) (A) y n° 2 (N/+28375,N/-28702) (B). T: amplificación realizada en ausencia de ARN. MW: marcador de ADN.
- la figura 12 ilustra la amplificación por RT-PCR en tiempo real de ARN sintético del gen N del SRAS-CoV: unas cantidades decrecientes de ARN sintético en replicado (repli. ; pistas 16 a 29) así como el ARN viral diluido al 1/20x10⁻⁴ (Pista 32) se amplificaron mediante RT-PCR en tiempo real con la ayuda del kit "Light Cycler RNA Amplification Kit Hybridization Probes" y de los pares de cebadores y de sondas de la serie n° 2, en las condiciones descritas en el ejemplo 8.

- la figura 13 (figura 13.1 a 13.70) representa el mapa de restricción de la secuencia SEC ID nº 1 que corresponde al equivalente ADN del genoma de la cepa de SRAS-CoV procedente de la extracción catalogada bajo el número 031589.
- la figura 14 muestra el resultado del ensayo de serología SRAS por ELISA N indirecto (1ª serie de sueros ensayados)
- la figura 15 muestra el resultado del ensayo de serología SRAS por ELISA N indirecto (2º serie de sueros ensayados)
- la figura 16 presenta el resultado del ensayo de serología SRAS por ELISA N doble epítipo (1ª serie de sueros ensayados)
- la figura 17 muestra el resultado del ensayo de serología SRAS por ELISA N doble epítipo (2ª serie de sueros ensayados)
- la figura 18 ilustra el ensayo de reactividad de los anticuerpos monoclonales anti-N por ELISA sobre la nucleoproteína N nativa del SRAS-CoV. Los anticuerpos se ensayaron en forma de sobrenadantes de cultivo de hibridomas por un ELISA indirecto, utilizando un lisado irradiado de células VeroE6 infectadas por el SRAS-CoV como antígeno (curvas lisado SRAS). Un control negativo de reactividad se realiza para cada anticuerpo sobre un lisado de células VeroE6 no infectadas (curvas lisado negativo). Varios anticuerpos monoclonales de especificidad conocida se usaron como anticuerpos control negativos: para 1-3 dirigido contra los antígenos de los virus parainfluenza de tipo 1-3 (Bio-Rad) y gripe B dirigido contra los antígenos del virus de la gripe de tipo B (Bio-Rad).
- la figura 19 ilustra el ensayo de reactividad de los anticuerpos monoclonales anti-N del SRAS-CoV por ELISA sobre los antígenos nativos de coronavirus humano 229E (HCoV-229E). Los anticuerpos se ensayaron en forma de sobrenadantes de cultivo de hibridomas por un ensayo ELISA indirecto, utilizando un lisado de células MRC-5 infectadas por el coronavirus humano 229E como antígeno (curvas lisado 229E). Un control negativo de inmunoreactividad se realiza, para cada anticuerpo, sobre un lisado de células MRC-5 no infectadas (curvas lisado negativo). El anticuerpo monoclonal 5-11H.6 dirigido contra la proteína S del coronavirus humano 229E (Sizun y otros 1998, J. Virol. Met. 72: 145-152) se utiliza como anticuerpo control positivo. Los anticuerpos para 1-3 dirigidos contra los antígenos de los virus parainfluenza de tipo 1-3 (Bio-Rad) y gripe B dirigido contra los antígenos del virus de la gripe de tipo B (Bio-Rad) se añadieron al panel de los anticuerpos monoclonales ensayados.
- la figura 20 muestra un ensayo de reactividad de los anticuerpos monoclonales anti-N del SRAS-CoV por transferencia western sobre la nucleoproteína N nativa del SRAS-CoV desnaturalizado. Un lisado de células VeroE6 infectadas por el SRAS-CoV- se preparó en el tampón de depósito según Laemmli y puesto a migrar en un gel SDS al 12% de poli(acrilamida), después las proteínas se transfirieron sobre una membrana de PVDF. Los anticuerpos monoclonales anti-N ensayados se utilizaron para el inmunoensayo a la concentración de 0,05 µg/ml. El revelado se hace con unos anticuerpos anti-IgG (H+L) de ratón acoplados a la peroxidasa (NA93IV, Amersham) y el sistema ECL+. Dos anticuerpos monoclonales se usaron como controles negativos de reactividad: gripe B dirigido contra los antígenos del virus de la gripe de tipo B (Bio-Rad) y para 1-3 dirigido contra los antígenos de los virus parainfluenza de tipo 1-3 (Bio-Rad).
- la figura 21 presenta los plásmidos de expresión en células de mamíferos de la proteína S del SRAS-CoV. El ADNc de la S del SRAS-CoV se insertó entre los sitios BamH1 y Xho1 del plásmido de expresión ACNpc3.1(+) (Clontech) para obtener el plásmido ADNpc-S y entre los sitios Nhe1 y Xho1 del plásmido de expresión pCI (Promega) para obtener el plásmido pCI-S- Las secuencias WPRE y CTE se insertaron en cada uno de los dos plásmidos ADNpc-S y pCI-S entre los sitios Xho1 y Xba1 para obtener respectivamente los plásmidos ADNpc-S-CTE, ADNpc-S-WPRE, pCI-S-CTE y pCI-S-WPRE.

SP: péptido señal predicho (aa 1-3) con el programa signalP v2.0 (Nielsen y otros, 1997, Protein Engineering, 10: 1-6)

TM: región transmembranaria predicha (aa 1196-1218) con el programa TMHMM v2.0 (Sonnhammer y otros, 1998, Proc. of Sixth Int. Conf. on Intelligent Systems for Molecular Biology, p 175-182, AAAI Press). Se debe señalar que los aminoácidos W1194 y P1195 forman parte posiblemente de la región transmembranaria con probabilidades respectivas de 0,13 y 0,42

P-CMV: promotor inmediato/precoz del citomegalovirus. BGH pA: señal de poliadenilación del gen de la hormona de crecimiento bovina

SV40 late pA: señal de poliadenilación tardía del virus SV40

SD/SA: sitios donante y receptor de empalme

WPRE: secuencias del "Woodchuck Hepatitis Virus posttranscriptional regulatory element" del virus de la hepatitis de la marmota

CTE: secuencias del "constitutive transport element" del retrovirus símico de Mason-Pfizer

- la figura 22 ilustra la expresión de la proteína S después de la transfección de células VeroE6. Se prepararon extractos celulares 48 horas después de la transfección de células VeroE6 por los plásmidos ADNpc, ADNpc-s, pCI y pCI-S. Extractos celulares se prepararon también 18 horas después de la infección por el virus recombinante de la vacuna antivariólica VV-TF7.3 y transfección por los plásmidos ADNpc o ADNpc-S. A título de control, unos extractos de células VeroE6 se prepararon 8 horas después de la infección por el SRAS-CoV con una multiplicidad de infección de 3. Se separaron sobre un gel SDS al 8% de acrilamida y fueron analizados mediante transferencia western con la ayuda de un anticuerpo policlonal de conejo anti-S y de un anticuerpo policlonal anti-IgG (H+L) de conejo acoplado con la peroxidasa (NA934V, Amersham). Una escala de masa molecular (kDa) se muestra en la figura.

SRAS-CoV: extracto de células VeroE6 infectadas por el SRAS-CoV

Mock: extracto control de células no infectadas

- la figura 23 ilustra el efecto de las secuencias CTE y WPRE sobre la expresión de la proteína S después de la transfección de células VeroE6 y 293T. Unos extractos celulares se prepararon 48 horas después de la transfección de células VeroE6 (A) o 293T (B) por los plásmidos ADNpc, ADNpc-S, ADNpc-S-CTE, ADNpc-S-WPRE, pCI-S, pCI-S-CTE y pCI-S-WPRE, separados sobre un gel SDS al 8% de acrilamida y analizados mediante transferencia western con la ayuda de un anticuerpo policlonal de conejo anti-S y de un anticuerpo policlonal anti-IgG (H+L) de conejo acoplado con la peroxidasa (NA934V, Amersham). Una escala de masa molecular (kDa) se lleva sobre la figura.

SRAS-CoV: extracto de células VeroE6 preparadas 8 horas después de la infección por el SRAS-CoV con una multiplicidad de infección de 3.

Mock: extracto control de células VeroE6 no infectadas

- la figura 24 presenta vectores lentivirales defectivos con "ADN flap" central para la expresión de la S du SRAS-CoV. El ADNc de la proteína S del SRAS-CoV se clonó en forma de un fragmento BamH1-Xho1 en el plásmido pTRIPAU3-CMV que contiene un vector lentiviral defectivo TRIP con "ADN flap" central (Sirven y otros, 2001, Mol. Ther., 3: 438-448) para obtener el plásmido pTRIP-S. Los casetes de expresión óptimos constituidos por el promotor inmediato/precoz del virus CMV, de una señal de empalme, del ADNc de la S y de una u otra de las señales post-transcripcionales CTE o WPRE se han sustituido al casete EF1a-EGFP del vector de expresión lentiviral defectivo de ADN FLAP central TRIPAU3-EF1 α (Sirven y otros, 2001, Mol. Ther., 3:438-448) para obtener los plásmidos pTRIP-SD/SA-S-CTE y pTRIP-SD/SA-S-WPRE.

SP: péptido señal

TM: región transmembranaria

P-CMV: promotor inmediato/precoz el citomegalovirus

P-EF1a: promotor del gen EF1 α

SD/SA: sitios donante y receptor de empalme

WPRES: secuencias del "Woodchuck Hepatitis Virus posttranscriptional regulatory element" del virus de la hepatitis de la marmota

CTE: secuencias del "constitutive transport element" del retrovirus símico de Mason-Pfizer

LTR: «Long terminal repeat»

Δ U3: LTR eliminado de las secuencias «promoter/enhancer»

cPPT: «polypurine tract cis-active sequence»

CTS: «central termination sequence»

- la figura 25 muestra el análisis mediante transferencia western de la expresión de la S du SRAS-CoV por líneas celulares transducidas por los vectores lentivirales TRIP-SD/SA-S-WPRE y TRIP-SD/SA-S-CTE. Se han preparado unos extractos celulares a partir de líneas FrhK4-S-CTE y FrhK4-S-WPRE establecidas después de la transducción por los vectores lentivirales TRIP-SD/SA-S-CTE y TRIP-SD /SA-S-WPRE respectivamente. Se han separado sobre un gel SDS al 8% de acrilamida y analizados mediante transferencia western con la ayuda de un anticuerpo policlonal de conejo anti-S y de un conjugado anti-IgG(H+L) de conejo acoplado con la peroxidasa. Una escala de masa molecular (kDa) se muestra en la figura.

T-: Extracto control de células FrhK-4

T+: Extracto de células FrhK-4 preparadas 24h después de la infección por el SRAS-CoV con una multiplicidad de infección de 3.

- la figura 26 se refiere al análisis de la expresión de polipéptido Ssol por líneas celulares transducidas por los vectores lentivirales TRIP-SD/SA-Ssol-WPRE y TRIP-SD/SA-Ssol-CTE. La secreción del polipéptido Ssol se ha buscado en el sobrenadante de una serie de clones celulares aislados después de la transducción de células FRhK-4 por los vectores lentivirales TRIP-SD/SA-Ssol-WPRE y TRIP-SD/SA-Ssol-CTE. 5 µl de sobrenadante diluidos al 1/2 en un tampón de depósito según Laemmli se analizaron mediante transferencia western revelado por un anticuerpo monoclonal anti-FLAG (M2, Sigma) y un conjugado anti-IgG(H+L) de ratón acoplado a la peroxidasa. T-: sobrenadante de la línea FRhK-4 parental. T+: sobrenadante de células BHK infectadas por un virus recombinante de la vacuna antivariólica que expresa el polipéptido Ssol. La flecha llena indica el polipéptido Ssol, mientras que la flecha hueca indica una reacción cruzada con una proteína de origen celular.

- la figura 27 muestra los resultados relativos al análisis del polipéptido Ssol purificado

A. 8, 2, 0,5 y 0,125 µg de polipéptido recombinante Ssol purificado mediante cromatografía de afinidad anti-FLAG y gel de filtración (G75) se separaron sobre gel SDS al 8% de poli(acrilamida). El polipéptido Ssol así como unas cantidades variables de marcadores de masa molecular (MM) se revelaron mediante coloración con nitrato de plata (Gelcode SilverSNAP stain kit II, Pierce).

B. Marcadores estándar para el análisis por espectrometría de masas SELDI-TOF

IgG: IgG bovino de MM 147300

ConA: conalbúmina de MM 77490

HRP: peroxidasa de rábano picante a título de control y de MM 43240

C. Análisis por espectrometría de masas (SELDI-TOF) del polipéptido recombinante Ssol. Los picos A y B corresponden al polipéptido Ssol simple y doblemente cargado.

D. Secuenciación del extremo N-terminal del polipéptido recombinante Ssol. Se realizaron 5 ciclos de degradación de Edman en fase líquida sobre un secuenciador ABI494 (Applied Biosystems).

- la figura 28 ilustra la influencia de una señal de corte y empalme y de las secuencias CTE y WPRE sobre la eficacia de la inmunización génica con la ayuda de ADN plasmídico que codifica la S du SRAS-CoV

A. Unos grupos de 7 ratones BALB/c se inmunizaron dos veces con 4 semanas de intervalo con la ayuda de 50 µg de ADN plasmídico de pCI, ADNpc-S, pCI-S, ADNpc-N y pCI-HA.

B. Unos grupos de 6 ratones BALB/c se inmunizaron dos veces con 4 semanas de intervalo con la ayuda de 2 µg, 10 µg ó 50 µg de ADN plasmídico de pCI, pCI-S, pCI-S-CTE y pCI-S-WPRE.

Los sueros inmunes extraídos 3 semanas después de la segunda inmunización se analizaron mediante ELISA indirecta utilizando un lisado de células VeroE6 infectadas por el SRAS-CoV como antígeno. Los títulos en anticuerpos anti-SRAS-CoV se calculan como la inversa de la dilución que produce una DO específica de 0,5 después del revelado por un anticuerpo policlonal anti-IgG de ratón acoplado a la peroxidasa (NA931V, Amersham) y de TMB (KPL).

- la figura 29 muestra la seroneutralización de la infectividad del SRAS-CoV por los anticuerpos inducidos en el ratón después de la inmunización génica con la ayuda de ADN plasmídico que codifica la S du SRAS-CoV. Unos grupos de sueros inmunes extraídos 3 semanas después de la segunda inmunización se realizaron para cada uno de los grupos de los experimentos descritos en la figura 28, y evaluados por su capacidad de seroneutralizar la infectividad de 100 TCID₅₀ de SRAS-CoV sobre células FRhK-4. Se realizan 4 puntos para cada una de las diluciones de razón 2 ensayadas a partir del 1/20. El título seroneutralizante se calcula según el método de Reed y Munsch como la inversa de la dilución que neutraliza la infectividad de 2 pozos sobre 4.

- A. Grupos de ratones BALB/c inmunizados dos veces con 4 semanas de intervalo con la ayuda de 50 µg de ADN plasmídico de pCI, ADNpc-S, pCI-S, ADNpc-N y pCI-HA. □: suero preinmune. ■: suero inmune.
- B. Grupos de ratones BALB/c inmunizados dos veces con 4 semanas de intervalo con la ayuda de 2 µg, 10 µg ó 50 µg de ADN plasmídico de pCI, pCI-S, pCI-S-CTE y pCI-S-WPRE.
- la figura 30 ilustra la inmunorreactividad del polipéptido recombinante Ssol frente a sueros de pacientes que padecen SRAS. La reactividad de sueros de pacientes se ha analizado mediante el ensayo ELISA indirecto contra unas fases sólidas preparadas con la ayuda del polipéptido recombinante Ssol purificado. Los anticuerpos de pacientes que reaccionan con la fase sólida a una dilución del 1/400 se revelan mediante un anticuerpo policlonal anti-IgG(H+L) humano acoplado a la peroxidasa (Amersham NA933V) y de TMB más H₂O₂ (KPL). Los sueros de casos posibles de SRAS son identificados por un número de orden del Centre National de Référence des virus influenzae así como por las iniciales del paciente y el número de días transcurridos desde el principio de los síntomas, llegado el caso. Los sueros TV son unos sueros controles de sujetos que han sido extraídos en Francia antes de la epidemia de SRAS aparecida en 2003.
 - la figura 31 muestra la inducción de anticuerpos dirigidos contra el SRAS-CoV después de la inmunización con el polipéptido recombinante Ssol. Dos grupos de 6 ratones se inmunizaron con 3 semanas de intervalo con 10 µg de polipéptido recombinante Ssol (grupo Ssol) adyuvado con el hidróxido de aluminio o, a título de control, del adyuvante sólo (grupo mock). Se realizaron tres inmunizaciones sucesivas y los sueros inmunes se extrajeron 3 semanas después de cada una de las inmunizaciones (IS1, IS2, IS3). Los sueros inmunes se analizaron por grupo para cada uno de los 2 grupos por ELISA indirecto utilizando un lisado de células VeroE6 infectadas por el SRAS-CoV como antígeno. Los títulos en anticuerpos anti-SRAS-CoV se calculan como la inversa de la dilución que produce una DO específica de 0,5 después del revelado por un anticuerpo policlonal anti-IgG de ratón acoplado a la peroxidasa (Amersham) y de TMB (KPL).
 - La figura 32 presenta la alineación nucleotídica de las secuencias del gen sintético 040530 con la secuencia del gen salvaje del aislado 031589 del SRAS-CoV. I-3059 corresponde a los nucleótidos 21406-25348 del aislado 031589 del SRAS-CoV depositado a la C.N.C.M. bajo el número 1-3059 (SEC ID nº 4, plásmido pSRAS-S). S-040530 es la secuencia del gen sintético 040530.
 - la figura 33 ilustra el uso de un gen sintético para la expresión de la S del SRAS-CoV. Unos extractos celulares preparados 48 horas después de la transfección de células VeroE6 (A) o 293T (B) por los plásmidos pCI, pCI-S, pCI-S-CTE, pCI-S-WPRE y pCI-Ssynth se prepararon sobre un gel SDS al 8% de acrilamida y analizados mediante transferencia western con la ayuda de un anticuerpo policlonal de conejo anti-S y de un anticuerpo policlonal anti-IgG(H+L) de conejo acoplado con la peroxidasa (NA934V, Amersham). La transferencia western se reveló mediante luminescencia (ECL+, Amersham) y adquisición sobre un dispositivo de creación de imagen numérica (FluorS, BioRad). Los niveles de expresión de la proteína S se midieron cuantificando las 2 bandas mayoritarias localizadas sobre la imagen.
 - la figura 34 presenta un esquema de la construcción de los virus de vacuna antivariólica recombinantes VV-TG-S, VV-TG-Ssol, VV-TN- S y VV-TN-Ssol
- A. Los ADNc de la proteína S y del polipéptido Ssol del SRAS-CoV se insertaron entre los sitios BamH1 y SmaI del plásmido de transferencia pTG186 para obtener los plásmidos pTG-S y pTG-Ssol.
- B. Las secuencias del promotor sintético 480 se sustituyeron después de las del promotor 7,5 mediante intercambio del fragmento Nde1-Pst1 de los plásmidos pTG186poli, pTG-S y pTG-Ssol para obtener los plásmidos de transferencia pTN480, pTN-S y pTN-Ssol.
- C. Secuencia del promotor sintético 480 tal como el contenido entre los sitios Nde1 y Pst1 de los plásmidos de transferencia de la serie pTN. Un sitio Asc1 se insertó para facilitar las manipulaciones ulteriores. Los sitios de restricción, así como la secuencia del promotor, son subrayados
- D. Los virus recombinantes de la vacuna antivariólica se obtienen mediante doble recombinación homóloga *in vivo* entre el casete TK de los plásmidos de transferencia de las series pTG y pTN, y el gen TK de la cepa Copenhagen del virus de la vacuna antivariólica.

SP: péptido señal predicha (aa 1-13) con el programa signalP v2.0 (Nielsen y otros, 1997, Protein Engineering, 10: 1-6)

TM: región transmembranaria predicha (aa 1196-1218) con el programa TMHMM v2.0 (Sonnhammer y otros, 1998, Proc. of Sixth Int. Conf. on Intelligent Systems for Molecular Biology, p 175-182, AAAI Press). Se debe de señalar que los aminoácidos W1194 y P1195 pertenecen posiblemente a la región transmembranaria con unas probabilidades respectivas de 0,13 y 0,42.

TK-L, TK-R: partes izquierda y derecha del gen de la timidinaquinasa del virus de la vacuna antivariólica

MCS: sitio múltiple de clonación

PE: promotor precoz

PL: promotor tardío

PL synth: promotor tardío sintético 480

- la figura 35 ilustra la expresión de la proteína S por los virus de la vacuna antivariólica recombinantes, analizada mediante transferencia western. Se han preparado unos extractos celulares 18 horas después de la infección de células CV1 por los virus de la vacuna antivariólica recombinantes VV-TG, VV-TG-S y VV-TN-S a una M.O.I. de 2 (A). A título de control, unos extractos de célula VeroE6 se prepararon 8 horas después de la infección por el SRAS-CoV con una multiplicidad de infección de 2. Unos extractos celulares se prepararon también 18 horas después de la infección de células CV1 por los virus de la vacuna antivariólica recombinantes VV-TG-S, VV-TG-Ssol, VV-TN, VV-TN-S y VV-TN-Ssol (B). Se prepararon sobre geles SDS al 8% de acrilamida y analizados mediante transferencia western con la ayuda de un anticuerpo policlonal de conejo anti-S y de un anticuerpo policlonal anti-IgG(H+L) de conejo acoplado a la peroxidasa (NA934V, Amersham). «1 µl» y «10 µl» indican las cantidades de extractos celulares depositadas sobre el gel. Una escala de masa molecular (kDa) se muestra en la figura.

SRAS-CoV: Extracto de células VeroE6 infectadas por el SRAS-CoV

Mock: Extracto control de células no infectadas

- la figura 36 muestra el resultado de un análisis mediante transferencia western de la secreción del polipéptido Ssol por los virus de la vacuna antivariólica recombinantes.

A. Unos sobrenadantes de células CV1 infectadas por el virus de la vacuna antivariólica recombinante VV-TN, diferentes clones de virus VV-TN-Ssol y por los virus VV-TG-Ssol o VV-TN-Sflag se recogieron 18 horas después de la infección de células CV1 a una M.O.I. de 2.

B. Unos sobrenadantes de células 293T, FRhK-4, BHK-21 y CV1 infectadas por duplicados (1,2) por el virus de la vacuna antivariólica recombinante VV-TN-Ssol a una M.O.I. de 2 se recogieron 18 horas después de la infección. El sobrenadante de células CV1 infectadas por el virus VV-TN se recogió también a título de control (M).

Todos los sobrenadantes se han separado sobre gel SDS a 8% de acrilamida según Laemmli, y analizados mediante transferencia western con la ayuda de un anticuerpo monoclonal de ratón anti-FLAG y de un anticuerpo policlonal anti-IgG(H+L) de ratón acoplado a la peroxidasa (NA931V, Amersham) (A) o con la ayuda de un anticuerpo policlonal de conejo anti-S y de un anticuerpo policlonal anti-IgG(H+L) de conejo acoplado con la peroxidasa (NA934V, Amersham) (B). Una escala de masa molecular (kDa) se muestra en la figura.

- la figura 37 muestra el análisis del polipéptido Ssol, purificado mediante gel SDS de poli(acrilamida

10, 5 y 2 µl de polipéptido recombinante Ssol purificado mediante cromatografía de afinidad anti-FLAG se separaron sobre gel SDS en gradiente del 4 al 15% de poli(acrilamida. El polipéptido Ssol, así como unas cantidades variables de marcadores de masa molecular (MM), se revelaron mediante coloración con nitrato de plata (Gelcode SilverSNAP stain kit II, Pierce).

- la figura 38 ilustra la inmunorreactividad del polipéptido recombinante Ssol producido por el virus de la vacuna antivariólica recombinante VV-TN-Ssol frente a sueros de pacientes que padecen SRAS. La reactividad de sueros de pacientes se analizó mediante un ensayo de ELISA indirecto contra unas fases sólidas preparadas con la ayuda del polipéptido recombinante Ssol purificado. Los anticuerpos de pacientes que reaccionan con la fase sólida a una dilución del 1/100 y 1/400 se revelan mediante un anticuerpo policlonal anti-IgG(H+L) humano acoplado a la peroxidasa (Amersham NA933V) y de TMB más H₂O₂ (KPL). Los sueros de casos posibles de SRAS son identificados mediante un número del tipo del Centre National de Référence des virus influenzae, así como por las iniciales del paciente y el número de días transcurridos desde el principio de los síntomas, llegado el caso. Los sueros TV son unos sueros controles de sujetos que fueron extraídos en Francia antes de la epidemia de SRAS aparecida en 2003.
- la figura 39 muestra la respuesta en anticuerpo anti-SRAS-CoV en el ratón después de la inmunización por los virus de la vacuna antivariólica recombinantes. Se inmunizaron unos grupos de 7 ratones BALB/c por vía i.v.

dos veces con 4 semanas de intervalo por 106 u.f.p. de virus de la vacuna antivariólica recombinantes VV-TG, VV-TG-HA, VV-TG-S, VV-TG-Ssol, VV-TN, VV-TN-S, VV-TN-Ssol.

A. Se realizaron unos grupos de sueros inmunes extraídos 3 semanas después de cada una de las dos inmunizaciones para cada uno de los grupos, y se analizaron mediante ELISA indirecto utilizando un lisado de células VeroE6 infectadas por el SRAS-CoV como antígeno. Los títulos en anticuerpos anti-SRAS-CoV son calculados como la inversa de la dilución que produce una DO específica de 0,5 después del revelado por un anticuerpo policlonal anti-IgG de ratón acoplado a la peroxidasa (NA931V, Amersham) y de TMB (KPL).

B. Los grupos de suero inmunes se evaluaron para su capacidad para seroneutralizar la infectividad de 100 TCID₅₀ del SRAS-CoV sobre células FRhK-4. Se realizan 4 puntos para cada una de las diluciones 2 veces ensayadas a partir del 1/20. El título seroneutralizante se calcula según el método de Reed y Munsch como la inversa de la dilución que neutraliza la infectividad de 2 pozos sobre 4.

- la figura 40 describe la construcción de los virus recombinantes MVSchw2-SRAS-S y MVSchw2-SRAS-Ssol.

A. El vector rubeola es un genoma completo de la cepa vacínea Schwarz del virus de la rubeola (MV) en el que se ha introducido una unidad de transcripción suplementaria (Combredet, 2003, Journal of Virology, 77: 11546-11554). La expresión de las fases abiertas de lectura (ORF) suplementarias está controlada por los elementos que actúan en cis necesarias para la transcripción, para la formación del capuchón y para la poliadenilación del transgen, que se copiaron de los elementos presentes en la unión N/P. 2 vectores diferentes permiten la inserción entre los genes P (fosfoproteína) y M (matriz) por una parte y H (hemaglutinina) y L (polimerasa) por otra parte.

B. Los genomas recombinantes MVSchw2-SRAS-S y MVSchw2-SRAS-Ssol del virus de la rubeola se construyeron mediante la inserción de los ORF de la proteína S y del polipéptido Ssol dentro de una unidad de transcripción suplementaria localizada entre los genes P y M del vector.

Los diferentes genes del virus de la rubeola (MV) son indicados: N (nucleoproteína), PVC (fosfoproteína y proteína V/C), M (matriz), F (fusión), H (hemaglutinina), L (polimerasa). T7 = promotor del ARN polimerasa T7, hh = ribozima "hammerhead", T7t = secuencia terminadora de la ARN polimerasa del fago T7, ∂ = ribozima del virus de la hepatitis ∂ , (2), (3) = unidades de transcripción suplementarias (ATU).

Tamaño del genoma de MV: 15894 nt.

SP: péptido señal

TM: región transmembranaria

FLAG: etiqueta FLAG

- la figura 41 ilustra la expresión de la proteína S por los virus de la rubeola recombinantes, analizada mediante transferencia western. Se prepararon extractos citoplásmicos después de la infección de células Vero por diferentes pasos de los virus MVSchw2-SRAS-S y MVSchw2-SRAS-Ssol y el virus salvaje MWSchw a título de control. Unos extractos celulares en tampón de depósito, según Laemmli, se prepararon también 8 horas después de la infección de células VeroE6 por el SRAS-CoV, con una multiplicidad de infección de 3. Se han preparado sobre un gel SDS al 8% de acrilamida y analizados mediante transferencia western con la ayuda de un anticuerpo policlonal de conejo anti-S y de un anticuerpo policlonal anti-IgG(H+L) de conejo acoplado a la peroxidasa (NA934V, Amersham).

Una escala de masa molecular (kDa) se muestra en la figura.

Pn: enésimo paso del virus después del cocultivo de células 293-3-46 y Vero.

SRAS-CoV: extracto de células VeroE6 infectadas por el SRAS-CoV

Mock: extracto control de células VeroE6 no infectadas

- la figura 42 muestra la expresión de la proteína S por los virus de la rubeola recombinantes, analizada mediante inmunofluorescencia

Se infectaron células Vero en monocapas sobre láminas de vidrio mediante el virus salvaje MWSchw (A) o los virus MVSchw2-SRAS-S (B) y MVSchw2-SRAS-Ssol (C). Cuando los sincitios alcanzaron del 30 al 40% de confluencia (A., B.) o 90-100% (C), las células se fijaron, permeabilizaron y marcaron mediante anticuerpos policlonales de conejo anti-SRAS-CoV y un conjugado anti-IgG(H+L) de conejo acoplado al FITC (Jackson).

- la figura 43 ilustra el análisis mediante transferencia western de la inmunorreactividad de sueros de conejo dirigidos contra los péptidos E1-12, E53-76 y M2-14. El conejo 20047 se inmunizó con el péptido E1-12 acoplado a la KLH. Los conejos 22234 y 22240 se inmunizaron con el péptido E53-76 acoplado a la KLH. Los conejos 20013 y 20080 se inmunizaron con el péptido M2-14 acoplado a la KLH. Los inmunosueros se analizaron mediante transferencia western con la ayuda de extractos de células infectadas por el SRAS-CoV (B) o con la ayuda de extractos de células infectadas por un virus recombinante de la vacuna antivariólica que expresa la proteína E (A) o M (C) del aislado 031589 del SRAS-CoV. Las inmunohuellas se recogieron con la ayuda de un conjugado anti-IgG(H+L) de conejo acoplado a la peroxidasa (NA934V, Amersham).

La posición de las proteínas E y M se indica mediante una flecha.
Una escala de masa molecular (kDa) se muestra en la figura.

Sin embargo, se debe entender, por supuesto, que estos ejemplos son dados únicamente a título de ilustración del objeto de la invención, y no constituyen de ninguna manera una limitación.

Ejemplo 1: Clonación y secuenciación del genoma de la cepa de SRAS-CoV procedente de la extracción catalogada bajo el número 031589

El ARN de la cepa de SRAS-CoV se ha extraído a partir de la extracción de lavado broncoalveolar catalogado bajo el número 031589, efectuado sobre un paciente del hospital francés de Hanoi (Vietnam) que padece SRAS.

El ARN aislado se utilizó como plantilla para amplificar los ADNc que corresponden a los diferentes cuadros de lectura abiertos del genoma (ORF1a, ORF1b, ORF-S, ORF-E, ORF-M, ORF-N (incluyendo los ORF-13 y ORF-14), ORF3, ORF4, ORF7 a ORF11), y a los extremos 5' y 3' no codificantes. Las secuencias de los cebadores y de las sondas utilizadas para la amplificación/detección se han definido según la secuencia nucleotídica disponible del SRAS-CoV.

A continuación, los cebadores y las sondas son identificados mediante: la letra S, seguida de una letra que indica la región correspondiente del genoma (L para el extremo 5' que incluye ORF1a y ORF1b; S, M y N para los ORF-S, ORF-M, ORF-N, SE y MN para las regiones intergénicas correspondientes), después, eventualmente, de Fn, Rn, con n incluido entre 1 y 6 que corresponde a los cebadores utilizados para la PCR anidada o imbricada (par F1 + R1 para la primera amplificación, par F2 + R2 para la segunda amplificación, etc.), después de +/- o -/- que corresponde a un cebador sentido o antisentido y finalmente unas posiciones de los cebadores en referencia a la secuencia Genbank AY27411.3; para los cebadores S y N sentido y antisentido y los otros cebadores sentido únicamente, cuando una sola posición está indicada ésta corresponde a la del extremo 5' de una sonda o de un cebador de aproximadamente 20 bases; para los cebadores antisentido diferentes de los cebadores S y N, cuando se indica una sola posición, ésta corresponde a la del extremo 3' de una sonda o de un cebador de aproximadamente 20 bases.

Los productos de amplificaciones así generados se secuenciaron con la ayuda de cebadores específicos a fin de determinar la secuencia completa del genoma de la cepa de SRAS-CoV procedente de la extracción catalogada bajo el número 031589. Estos productos de amplificación, con la excepción de los que corresponden a los ORF1a y ORF1b, se clonaron después en unos vectores de expresión a fin de producir las proteínas virales correspondientes y los anticuerpos dirigidos contra estas proteínas, en particular mediante inmunización a base de ADN.

1. Extracción de los ARN

Los ARN han sido extraídos con la ayuda del kit *Qlamp viral RNA extraction mini* (QIAGEN) siguiendo las recomendaciones del fabricante. De manera más precisa: se han mezclado 140 µl de la extracción y 560 µl de tampón AVL vigorosamente durante 15 segundos, incubados durante 10 minutos a temperatura ambiente y después centrifugados brevemente a velocidad máxima. Se han añadido 560 µl de etanol al 100% al sobrenadante y la mezcla así obtenida se agitó muy vigorosamente durante 15 s. Después, se han depositado 630 µl de la mezcla en la columna.

La columna se colocó sobre un tubo de 2 ml, se centrifugó durante 1 minuto a 8000 rpm, y después el resto de la mezcla se depositó sobre la misma columna, y se centrifugó nuevamente durante 1 minuto a 8000 rpm, y se transfirió la columna sobre un tubo de 2 ml limpio. Después, se añadieron 500 µl de tampón AW1 sobre la columna, se centrifugó después la columna durante 1 minuto a 8000 rpm y se eliminó el eluido. Se añadieron 500 µl de tampón AW2 sobre la columna que se centrifugó después durante 3 minutos a 14000 rpm y se transfirió en un tubo de 1,5 ml. Finalmente, se añadieron 60 µl de tampón AVE sobre la columna que se incubó durante 1 a 2 minutos a temperatura ambiente y después se centrifugó durante 1 minuto a 8000 rpm. El eluido que corresponde al ARN purificado se recuperó y se congeló a -20°C.

2. Amplificación, secuenciación y clonación de los ADNc

2.1) ADNc que codifica la proteína S

Los ARN extraídos a partir de la extracción se sometieron a una transcripción inversa con la ayuda de oligonucleótidos hexaméricos de secuencia aleatoria (pdN6), a fin de producir unos fragmentos de ADNc.

La secuencia que codifica la glicoproteína S del SRAS-CoV se amplificó en forma de dos fragmentos de ADN superpuestos: fragmentos 5' (SRAS-Sa, SEC ID nº 5) y fragmento 3' (SRAS-Sb, SEC ID nº 6), realizando dos amplificaciones sucesivas con la ayuda de cebadores imbricados. Los amplicones así obtenidos se secuenciaron, clonaron en el vector plasmídico PCR 2.1-TOPO™ (IN VITROGEN), y después se determinó la secuencia de los ADNc clonados.

a) clonación y secuenciación de los fragmentos Sa y Sb

a₁) síntesis del ADNc

La mezcla de reacción que contiene: ARN (5 µl), H₂O ppi (3,5 µl), tampón de transcriptasa inversa 5X (4 µl), dNTP 5 mM (2 µl), pdN6 100 µg/ml (4 µl), RNasin 40 U/ul (0,5 µl) y transcriptasa inversa AMV-RT, 10 U/ul, PROMEGA (1 µl) se incubó en un termociclador en las condiciones siguientes: 45 min. a 42°C, 15 min. a 55°C, 5 min. a 95°C, y después se mantuvo el ADNc obtenido a +4°C.

a₂) primera amplificación PCR

Los extremos 5' y 3' del gen S se amplificaron respectivamente con los pares de cebadores S/F1+/21350-21372 y S/R1-/23518-23498, S/F3+/23258-23277 y S/R3-/25382-25363. La mezcla de reacción de 50 µl que contiene: ADNc (2 µl), cebadores 50 µM (0,5 µl), tampón 10 X (5 µl), dNTP 5 mM (2 µl), Taq Expand High Fidelity, Roche (0,75 µl) y H₂O (39, 75 µl) se amplificó en un termociclador, en las condiciones siguientes: una etapa inicial de desnaturalización a 94°C durante 2 minutos, seguida de 40 ciclos que comprenden: una etapa de desnaturalización a 94°C durante 30 s, una etapa de hibridación a 55°C durante 30 s y después una etapa de elongación a 72°C durante 2 minutos 30 s, con 10 s de elongación suplementaria a cada ciclo, y después una etapa final de elongación a 72°C durante 5 minutos.

a₃) segunda amplificación PCR

Los productos de la primera amplificación PCR (amplicones 5' y 3') han sufrido una segunda etapa de amplificación PCR (PCR anidada) en condiciones idénticas a las de la primera amplificación, con los pares de cebadores S/F2+/21406-21426 y S/R2-/23454-23435, y S/F4+/23322-23341 y S/R4-/25348-25329, respectivamente para el amplicón 5' y el amplicón 3'.

a₄) clonación y secuenciación de los fragmentos Sa y Sb

Los amplicones Sa (extremo 5') y Sb (extremo 3') así obtenidos se purificaron con la ayuda del kit *QIAquick PCR purification* (QIAGEN), siguiendo las recomendaciones del fabricante, después se clonaron en el vector PCR2.1-TOPO (kit Invitrogen), para dar los plásmidos denominados SRAS-S1 y SRAS-S2.

El ADN de los clones Sa y Sb se aisló y después se secuenció el inserto con la ayuda del kit Big Dye, Applied Biosystem® y de los cebadores universales M13 directo y M13 inverso, así como de los cebadores: S/S+/21867, S/S+/22353, S/S+/22811, S/S+/23754, S/S+/24207, S/S+/24699, S/S+/24348, S/S-/24209, S/S-/23630, S/S-/23038, S/S-/22454, S/S-/21815, S/S-/24784, S/S+/21556, S/S+/23130 y S/S+/24465, siguiendo las instrucciones del fabricante; las secuencias de los fragmentos Sa y Sb así obtenidas corresponden a las secuencias SEC ID nº 5 y SEC ID nº 6 en el listado de secuencias adjunto en anexo.

El plásmido denominado SRAS-S1 se depositó bajo el nº 1-3020, el 12 de mayo de 2003, en la Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15; contiene un fragmento 5' de la secuencia del gen S de la cepa de SRAS-CoV procedente de la extracción catalogada bajo el nº 031559, tal como se ha definido anteriormente, correspondiendo dicho fragmento denominado Sa a los nucleótidos de las posiciones 21406 a 23454 (SEC ID nº 5), en referencia a la secuencia Genbank AY274119.3 Tor2.

El plásmido denominado TOP10F'-SRAS-S2 se depositó bajo el nº 1-3019, el 12 de mayo de 2003, en la Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15; contiene un fragmento 3' de la secuencia del gen S de la cepa de SRAS-CoV procedente de la extracción catalogada bajo el nº 031589, tal como se ha definido anteriormente, correspondiendo dicho fragmento denominado Sb a los nucleótidos de las posiciones 23322 a 25348 (SEC ID nº 6), en referencia a la secuencia Genbank nº de acceso AY274119.3.

b) clonación y secuenciación del ADNc completo (clon SRAS-S de 4 kb)

El ADNc S completo se obtuvo a partir de los clones SRAS-S1 y SRAS-S2 antes citados, de la siguiente manera:

- 1) se realizó una reacción de amplificación PCR sobre un clon SRAS-S2 en presencia del cebador S/R4/-/25348-25329 antes citado y del cebador S/S/+24696-24715: se obtuvo un amplicón de 633 pb,
- 2) se realizó otra amplificación PCR sobre otro clon SRAS-S2, en presencia de los cebadores S/F4/+23322-23341 antes citados y S/S/-24803-24784: se obtuvo un amplicón de 1481 pb,

La reacción de amplificación se realizó en las condiciones tales como se han definido anteriormente para la amplificación de los fragmentos Sa y Sb, con la excepción de que se efectuaron 30 ciclos de amplificaciones que comprenden una etapa de desnaturalización a 94°C durante 20 s y una etapa de elongación a 72°C durante 2 minutos 30 s.

3) los 2 amplicones (633 pb y 1481 pb) se purificaron en las condiciones tales como se han definido anteriormente para los fragmentos Sa y Sb.

4) se realizó otra reacción de amplificación PCR con la ayuda de los cebadores S/F4/+23322-23341 y S/R4/-/25348-25329 antes citados, sobre los amplicones purificados obtenidos en 3). La reacción de amplificación se realizó en las condiciones tales como se han definido anteriormente para la amplificación de los fragmentos Sa y Sb, con la excepción de que se efectuaron 30 ciclos de amplificaciones.

El amplicón de 2026 pb así obtenido se purificó, se clonó en el vector PCR2.1-TOPO y después se secuenció como anteriormente, con la ayuda de los cebadores tales como se han definido anteriormente para los fragmentos Sa y Sb. El clon así obtenido se denominó clon 3'.

5) El clon SRAS-S1 obtenido anteriormente y el clon 3' se han digerido por *EcoR I*, se purificaron las bandas de aproximadamente 2 kb así obtenidas sobre gel y después se amplificaron mediante PCR con los cebadores S/F2/+21406-21426 y S/R4/-/25348-25329 antes citados. La reacción de amplificación se realizó en las condiciones tales como se han definido anteriormente para la amplificación de los fragmentos Sa y Sb, con la excepción de que se efectuaron 30 ciclos de amplificaciones. El amplicón de aproximadamente 4 kb se purificó y se secuenció. Después, se clonó en el vector PCR2.1-TOPO para dar el plásmido denominado SRAS-S, y se secuenció el inserto contenido en este plásmido como anteriormente, con la ayuda de los cebadores tales como se han definido anteriormente para los fragmentos Sa y Sb. Las secuencias de ADNc del inserto y del amplicón que codifica la proteína S, corresponden respectivamente a las secuencias SEC ID nº 4 y SEC ID nº 2 en el listado de secuencias adjunto en anexo, éstas codifican para la proteína S (SEC ID nº 3).

La secuencia del amplicón que corresponde al ADNc que codifica la proteína S de la cepa de SRAS-CoV procedente de la extracción nº031589 presenta las dos mutaciones siguientes con respecto a las secuencias que corresponden respectivamente a los aislados Tor2 y Urbani, siendo las posiciones de las mutaciones indicadas en referencia a la secuencia completa del genoma del aislado Tor2 (Genbank AY274119.3):

- g/t en posición 23220; el codón de alanina (gct) en posición 577 de la secuencia en aminoácidos de la proteína S de Tor2 está sustituido por un codón de serina (tct),
- c/t en posición 24872: esta mutación no modifica la secuencia en aminoácidos de la proteína S, y

El plásmido denominado SRAS-S se depositó bajo el nº 1-3059, el 20 de junio de 2003, en la Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15; contiene la secuencia de ADNc que codifica la proteína S de la cepa de SRAS-CoV procedente de la extracción catalogada bajo el nº 031589, secuencia que corresponde a los nucleótidos de las posiciones 21406 a 25348 (SEC ID nº 4), en referencia a la secuencia Genbank AY274119.3.

2.2) ADNc que codifica las proteínas M y E

Los ARN procedentes de la extracción 031589, extraídos como anteriormente, se sometieron a una transcripción inversa, asociada, durante la misma etapa (kit Titan One Step RT-PCR®, Roche), a una reacción de amplificación por PCR, con la ayuda de los pares de cebadores:

- S/E/F1/+26051-26070 y S/E/R1/-26455-26436 para amplificar el ORF-E, y
- S/M/F1/+26225-26244 y S/M/R1/-27148-27129 para amplificar el ORF-M.

Una primera mezcla de reacción que contiene: 8,6 µl de H₂O_{pp}i, 1 µl de dNTP (5 mM), 0,2 µl de cada uno de los cebadores (50 µM), 1,25 µl de DTT (100 mM) y 0,25 µl de ARNsin (40 UI/µl) se combinó con una segunda mezcla de

reacción que contiene: 1 µl de ARN, 7 µl de H₂O_{pp}i, 5 µl de tampón de RT-PCR 5X y 0,5 µl de mezcla de enzima, y se incubaron las mezclas combinadas en un termociclador en las condiciones siguientes: 30 min. a 42°C, 10 min. a 55°C, 2 min. a 94°C seguido de 40 ciclos que comprenden una etapa de desnaturalización a 94°C durante 10 s, una etapa de hibridación a 55°C durante 30 s y una etapa de elongación a 68°C durante 45 s, con 3 s de incremento por ciclo, y finalmente una etapa de elongación terminal a 68°C durante 7 min.

Los productos de amplificación así obtenidos (amplicones M y E) han sufrido una segunda amplificación PCR (PCR anidada) utilizando el kit Expand High-Fi[®], Roche, con la ayuda de los pares de cebadores:

- S/E/F2/+26082-26101 y S/E/R2/-26413-26394 para el amplicón E, y
- S/M/F2/+26330-26350 y S/M/R2/-27098-27078 para el amplicón M.

Conteniendo la mezcla de reacción: 2 µl del producto del primer PCR, 39,25 µl de H₂O_{pp}i, 5 µl de tampón 10X que contiene MgCl₂, 2 µl de dNTP (5 mM), 0,5 µl de cada uno de los cebadores (50 µM) y 0,75 µl de mezcla de enzima se incubaron en un termociclador en las condiciones siguientes: una etapa de desnaturalización a 94°C durante 2 min. seguido de 30 ciclos que comprenden una etapa de desnaturalización a 94°C durante 15 s, una etapa de hibridación a 60°C durante 30 s y una etapa de elongación a 72°C durante 45 s, con 3 s de incremento por ciclo, y finalmente una etapa de elongación terminal a 72°C durante 7 min. Los productos de amplificación obtenidos que corresponden a los ADNc que codifican para las proteínas E y M se secuenciaron como anteriormente, con la ayuda de los cebadores: S/E/F2/+26082 y S/E/R2/-26394, S/M/F2/+26330, S/M/R2/-27078 antes citados y de los cebadores S/M/+26636-26655 y S/M/-26567-26548. Después, se clonaron, como anteriormente, para dar los plásmidos denominados SRAS-E y SRAS-M. El ADN de estos clones se aisló y se secuenció después con la ayuda de los cebadores universales M13 directo y M13 inverso, así como unos cebadores S/M/+26636 y S/M/-26548 antes citados.

La secuencia del amplicón que representa el ADNc que codifica la proteína E (SEC ID n° 13) de la cepa de SRAS-CoV procedente de la extracción n°031589, no comprende diferencias con respecto a las secuencias que corresponden a unos aislados AY274119.3-Tor2 y AY278741-Urbani. La secuencia de la proteína E de la cepa de SRAS-CoV 031589 corresponde a la secuencia s ID n° 14 en el listado de secuencias adjunto en anexo.

El plásmido denominado SRAS-E se depositó bajo el n° 1-3046, el 28 de mayo de 2003, en la Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15; contiene la secuencia de ADNc que codifica la proteína E de la cepa de SRAS-CoV procedente de la extracción catalogada bajo el n° 031589, tal como se ha definido anteriormente, secuencia que corresponde a los nucleótidos de las posiciones 26082 a 26413 (SEC ID n° 15), en referencia a la secuencia Genbank n° de acceso AY274119.3.

La secuencia del amplicón que representa el ADNc que codifica la M (SEC ID n° 16) de la cepa de SRAS-CoV procedente de la extracción n°031589 no comprende diferencias con referencia a la secuencia correspondiente del aislado AY274119.3-Tor2. Sin embargo, en la posición 26857, el aislado AY278741-Urbani comprende una c y la secuencia de la cepa de SRAS-CoV procedente de la extracción catalogada bajo el n°031589 una t. Esta mutación desemboca en una modificación de la secuencia en aminoácidos de la proteína correspondiente: en la posición 154, una prolina (AY278741-Urbani) se cambia a serina en la cepa de SRAS-CoV procedente de la extracción catalogada bajo el n°031589. La secuencia de la proteína M de la cepa de SRAS-CoV procedente de la extracción catalogada bajo el n°031589 corresponde a la secuencia SEC ID n° 17 en el listado de secuencias adjunto en anexo.

El plásmido denominado SRAS-M se depositó bajo el n° 1-3047, el 28 de mayo de 2003, en la Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15; contiene la secuencia de ADNc que codifica la proteína M de la cepa de SRAS-CoV procedente de la extracción catalogada bajo el n° 031589, tal como se ha definido anteriormente; secuencia que corresponde a los nucleótidos de las posiciones 26330 a 27098 (SEC ID n° 18), en referencia a la secuencia Genbank n° de acceso AY274119.3.

2.3) ADNc que corresponde a los ORF3, ORF4, ORF7 a ORF11

La misma estrategia de amplificación, de clonación y de secuenciación, se utilizó para obtener los fragmentos de ADNc que corresponden respectivamente a los ORF siguientes: ORF 3, ORF4, ORF7, ORF8, ORF9, ORF10 y ORF11. Los pares de cebadores utilizados para la primera amplificación son:

- ORF3 y ORF4: S/SE/F1/+25069-25088 y S/SE/R1/-26300-26281
- ORF7 a ORF11: S/MN/F1/+26898-26917 y S/MN/R1/-28287-28266

Los pares de cebadores utilizados para la segunda amplificación son:

- ORF3 y ORF4: S/SE/F2/+25110-25129 y S/SE/R2/-26244-26225

- ORF7 a ORF11: S/MN/F2/+26977-26996 y S/MN/R2/-28218-28199

Las condiciones de la primera amplificación (RT-PCR) son las siguientes: 45 min. a 42°C, 10 min. a 55°C, 2 min. a 94°C, seguido de 40 ciclos que comprenden una etapa de desnaturalización a 94°C durante 15 s, una etapa de hibridación a 58°C durante 30 s y una etapa de elongación a 68°C durante 1 min., con 5 s de incremento por ciclo, y finalmente una etapa de elongación terminal a 68°C durante 7 min.

Las condiciones de la PCR anidada son las siguientes: una etapa de desnaturalización a 94°C durante 2 min. seguida de 40 ciclos que comprenden una etapa de desnaturalización a 94°C durante 20 s, una etapa de hibridación a 58°C durante 30 s y una etapa de elongación a 72°C durante 50 s, con 4 s de incremento por ciclo, y finalmente una etapa de elongación terminal a 72°C durante 7 min.

Los productos de amplificación obtenidos que corresponden a los ADNc que contienen respectivamente los ORF3 y 4 y los ORF7 a 11 se secuenciaron con la ayuda de los cebadores: S/SE+/25363, S/SE+/25835, S/SE-/25494, S/SE-/25875, S/MN/+/27839, S/MN+/27409, S/MN-/27836 S/MN-/27799 y se clonaron como anteriormente para los otros ORF, para dar los plásmidos denominados SRAS-SE y SRAS-MN. El ADN de estos clones se aisló y se secuenció con la ayuda de estos mismos cebadores, y unos cebadores universales M 13 sentido y M 13 anti-sentido.

La secuencia del amplicón que representa el ADNc de la región que contiene los ORF 3 y 4 (SEC ID nº 7) de la cepa de SRAS-CoV procedente de la extracción nº031589 comporta una diferencia nucleotídica con referencia a la secuencia correspondiente del aislado AY274119-Tor2. Esta mutación en la posición 25298 desemboca en una modificación de la secuencia en aminoácidos de la proteína correspondiente (ORF 3): en la posición 11, una arginina (AY274119-Tor2) se cambia a glicina en la cepa de SRAS-CoV procedente de la extracción nº031589. Sin embargo, no se ha identificado ninguna mutación con referencia a la secuencia correspondiente del aislado AY278741-Urbani. Las secuencias de los ORF 3 y 4 y la cepa de SRAS-CoV procedente de la extracción nº031589 corresponden respectivamente a las secuencias SEC ID nº 10 y 12 en el listado de secuencias adjunto en anexo.

El plásmido denominado SRAS-SE se depositó bajo el nº I-3126, el 13 de noviembre de 2003, en la Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15; contiene el ADNc que corresponde a la región situada entre el ORF-S y el ORF-E y que se superpone a ORF-E de la cepa de SRAS-CoV procedente de la extracción catalogada bajo el nº 031589, tal como se ha definido anteriormente, correspondiendo dicha región a los nucleótidos de las posiciones 25110 a 26244 (SEC ID nº 8), en referencia a la secuencia Genbank nº de acceso AY274119.3,

La secuencia del amplicón que representa el ADNc que corresponde a la región que contiene los ORF7 a ORF11 (SEC ID nº 19) de la cepa de SRAS-CoV procedente de la extracción nº031589 no comprende diferencias con respecto a las secuencias correspondientes de los aislados AY274119-Tor2 y AY278741-Urbani. Las secuencias de los ORF7 a 11 de la cepa de SRAS-CoV procedente de la extracción nº031589 corresponden respectivamente a las secuencias SEC ID nº 22, 24, 26, 28 y 30 en el listado de secuencias adjunto en anexo.

El plásmido denominado SRAS-MN se depositó bajo el nº I-3125, el 13 de noviembre de 2003, en la Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15; contiene la secuencia de ADNc que corresponde a la región situada entre el ORF-M y el ORF-N de la cepa de SRAS-CoV, procedente de la extracción catalogada bajo el nº 031589 y extraída en Hanoi, tal como se ha definido anteriormente, secuencia que corresponde a los nucleótidos de las posiciones 26977 a 28218 (SEC ID nº 20), en referencia a la secuencia Genbank nº de acceso AY274119.3.

La secuencia del amplicón que representa el ADNc que corresponde a la región que contiene los ORF7 a ORF11 (SEC ID nº 19) de la cepa de SRAS-CoV procedente de la extracción nº031589 no comprenden diferencias con respecto a las secuencias correspondientes de los aislados AY274119-Tor2 y AY278741-Urbani. Las secuencias de los ORF7 a 11 de la cepa de SRAS-CoV procedente de la extracción nº031589 corresponden respectivamente a las secuencias SEC ID nº 22, 24, 26, 28 y 30 en el listado de secuencias adjunto en anexo.

2.4) ADNc que codifica la proteína N y que incluye los ORF13 y ORF14

El ADNc se sintetizó y se amplificó como se describió anteriormente para los fragmentos Sa y Sb. De manera más precisa, la mezcla de reacción que contiene: 5 µl de ARN, 5 µl de H₂O ppi 4 µl de tampón de transcriptasa inversa 5X, 2 µl de dNTP (5 mM), 2 µl de oligo 20T (5 µM), 0,5 µl de RNasin (40 UI/µl) y 1, 5 µl de AMV-RT (10 UI/µl Promega) se incubó en un termociclador en las condiciones siguientes: 45 min. a 42°C, 15 min. a 55°C, 5 min. a 95°C, y después se mantuvo a +4°C.

Se realizó una primera amplificación PCR con el par de cebadores S/N/F3/+28023 y S/N/R3/-29480.

La mezcla de reacción como anteriormente para la amplificación de los fragmentos S1 y S2 se incubó en un termociclador, en las condiciones siguientes: una etapa inicial de desnaturalización a 94°C durante 2 min. seguida

de 40 ciclos que comprenden una etapa de desnaturalización a 94°C durante 20 s, una etapa de hibridación a 55°C durante 30 s y después una etapa de elongación a 72°C durante 1 min. 30 s con 10 s de elongación suplementaria en cada ciclo, y después una etapa final de elongación a 72°C durante 5 min.

El amplicón obtenido en la primera amplificación PCR ha sufrido una segunda etapa de amplificación PCR (PCR anidada) con el par de cebadores S/N/F4/+28054 y S/N/R4/-29430 en condiciones idénticas a las de la primera amplificación.

El producto de amplificación obtenido que corresponde al ADNc que codifica la proteína N de la cepa de SRAS-CoV procedente de la extracción n°031589 se secuenció con la ayuda de los cebadores: S/N/F4/+28054, S/N/R4/-29430, S/N/+28468, S/N/+28918 y S/N/-28607 y se clonó como anteriormente para los otros ORF, para dar el plásmido denominado SRAS-N. El ADN de estos clones se aisló y se secuenció con la ayuda de los cebadores universales M13 sentido y M13 anti-sentido, así como de los cebadores S/N/+28468, S/N/+28918 y S/N/-28607.

La secuencia del amplicón que representa el ADNc que corresponde a ORF-N y que incluye los ORF13 y ORF14 (SEC ID n° 36) de la cepa de SRAS-CoV procedente de la extracción n°031589 no comprenden diferencias con respecto a las secuencias correspondientes de los aislados AY274119.3-Tor2 y AY278741-Urbani. La secuencia de la proteína N de la cepa de SRAS-CoV procedente de la extracción n°031589 corresponde a la secuencia SEC ID n° 37 en el listado de secuencias adjunto en anexo.

Las secuencias de los ORF13 y 14 de la cepa de SRAS-CoV procedente de la extracción n°031589 corresponden respectivamente a las secuencias SEC ID n° 32 y 34 en el listado de secuencias adjunto en anexo.

El plásmido denominado SRAS-N se depositó bajo el n° I-3048, el 5 de junio de 2003, en la Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15; contiene el ADNc que codifica la proteína N de la cepa de SRAS-CoV procedente de la extracción catalogada bajo el n° 031589, tal como se ha definido anteriormente, secuencia que corresponde a los nucleótidos de las posiciones 28054 a 29430 (SEC ID n° 38), en referencia a la secuencia Genbank n° de acceso AY274119.3.

2.5) extremos 5' y 3' no codificantes

a) extremo 5' no codificante (5'NC)

a₁) síntesis del ADNc

Los ARN procedentes de la extracción 031589, extraídos como anteriormente, se sometieron a una transcripción inversa en las condiciones siguientes:

El ARN (15 µl) y el cebador S/L/-443 (3 µl de la concentración de 5 µM, se incubaron durante 10 min. a 75°C.

Después, se añadieron un tampón de transcriptasa inversa 5X (6 µl, INVITROGEN), unos dNTP 10 mM (1 µl), un DTT 0,1M (3 µl), y se incubó la mezcla a 50°C durante 3 min.

Finalmente, se añadió la transcriptasa inversa (3 µl de Superscript®, INVITROGEN) a la mezcla anterior, que se incubó a 50°C durante 1h30 y después a 90 °C durante 2 min.

El ADNc así obtenido se purificó con la ayuda del kit *QIAquick PCR purification* (QIAGEN), según las recomendaciones del fabricante.

b₁) Reacción a la Terminal Transferasa (TdT)

El ADNc (10 µl) se incubó durante 2 min. a 100°C, se conserva en hielo, y después se añaden: H₂O (2,5 µl), tampón TdT 5X (4 µl, AMERSHAM), dATP 5mM (2 µl) y TdT (1,5 µl, AMERSHAM). La mezcla así obtenida se incubó durante 45 min. a 37°C y después durante 2 min. a 65°C.

El producto obtenido se amplifica mediante una primera reacción PCR con la ayuda de los cebadores: S/L/-225-206 y anclaje 14T: 5'-AGATGAATTCGGTACCTTTTTTTTTTTTTTTT-3' (SEC ID n° 68). Las condiciones de la amplificación son las siguientes: una etapa inicial de desnaturalización a 94°C durante 2 min. es seguida de 10 ciclos que comprenden una etapa de desnaturalización a 94°C durante 10 s, una etapa de hibridación a 45°C durante 30 s y después una etapa de elongación a 72°C durante 30 s, y después de 30 ciclos que comprenden una etapa de desnaturalización a 94°C durante 10 s, una etapa de hibridación a 50°C durante 30 s y después una etapa de elongación a 72°C durante 30 s, y después de una etapa final de elongación a 72°C durante 5 min.

El producto de la primera amplificación PCR sufrió una segunda etapa de amplificación con la ayuda de los cebadores: S/L/-204-185 y anclaje 14T antes citado en condiciones idénticas a las de la primera amplificación. El

amplicón así obtenido se purificó, se secuenció con la ayuda del cebador S/L/-182-163 y después se clonó como anteriormente para los diferentes ORF, para dar el plásmido denominado SRAS-5'NC. El ADN de este clon se aisló y se secuenció con la ayuda de los cebadores universales M13 sentido y M13 anti-sentido y del cebador S/L/-182-163 antes citado.

El amplicón que representa el ADNc que corresponde al extremo 5'NC de la cepa de SRAS-CoV procedente de la extracción catalogada bajo el n° 031589 corresponde a la secuencia SEC ID n° 72 en el listado de secuencias adjunto en anexo; esta secuencia no comprende diferencias con respecto a las secuencias correspondientes de los aislados AY274119.3-Tor2 y AY278741-Urbani.

El plásmido denominado SRAS-5'NC se depositó bajo el n° I-3124, el 7 de noviembre de 2003, en la Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15; contiene el ADNc que corresponde al extremo 5' no codificante del genoma de la cepa de SRAS-CoV procedente de la extracción catalogada bajo el n° 031589, tal como se ha definido anteriormente, secuencia que corresponde a los nucleótidos de las posiciones 1 a 204 (SEC ID n° 39), en referencia a la secuencia Genbank n° de acceso AY274119.3.

b) extremo 3' no codificante (3'NC)

a₁) síntesis del ADNc

Los ARN procedentes de la extracción n° 031589, extraídos como anteriormente, se sometieron a una transcripción inversa, según el protocolo siguiente: conteniendo la mezcla de reacción: ARN (5 µl), H₂O (5 µl), tampón de transcriptasa inversa 5X (4 µl), dNTP 5 mM (2 µl), Oligo 20T 5 µM (2 µl), RNasin 40 U/µl (0,5 µl) y RT-AMV 10 UI/µl (1,5 µl, PROMEGA) se incubó en un termociclador, en las condiciones siguientes: 45 min. a 42°C, 15 min. a 55°C, 5 min. a 95°C, y después se mantuvo a +4°C.

El ADNc obtenido se amplificó mediante una primera reacción PCR con la ayuda de los cebadores S/N/+28468-28487 y anclaje 14T antes citado. Las condiciones de la amplificación son las siguientes: una etapa inicial de desnaturalización a 94°C durante 2 min. seguida de 10 ciclos que comprenden una etapa de desnaturalización a 94°C durante 20 s, una etapa de hibridación a 45°C durante 30 s y después una etapa de elongación a 72°C durante 50 s. y después de 30 ciclos que comprenden una etapa de desnaturalización a 94°C durante 20 s, una etapa de hibridación a 50°C durante 30 s y después una etapa de elongación a 72°C durante 50 s, y después de una etapa final de elongación a 72°C durante 5 min.

El producto de la primera amplificación PCR sufrió una segunda etapa de amplificación con la ayuda de los cebadores S/N/+28933-28952 y anclaje 14T antes citados, en condiciones idénticas a las de la primera amplificación. El amplicón así obtenido se purificó, se secuenció con la ayuda del cebador S/N/+29257-29278 y se clonó como anteriormente para los diferentes ORF, para dar el plásmido denominado SRAS-3'NC. El ADN de este clon se aisló y se secuenció con la ayuda de los cebadores universales M13 sentido y M13 anti-sentido y del cebador S/N/+29257-29278 antes citado.

El amplicón que representa el ADNc que corresponde al extremo 3'NC de la cepa de SRAS-CoV procedente de la extracción catalogada bajo el n° 031589 corresponde a la secuencia SEC ID n° 73 en el listado de secuencias adjunto en anexo; esta secuencia no comprende diferencias con respecto a las secuencias correspondientes de los aislados AY274119.3-Tor2 y AY278741-Urbani.

El plásmido denominado SRAS-3'NC se depositó bajo el n° I-3123, el 7 de noviembre de 2003, en la Collection Nationale de Cultures de Microorganismes, 25 rue du Docteur Roux, 75724 Paris Cedex 15; contiene la secuencia de ADNc que corresponde al extremo 3' no codificante del genoma de la cepa de SRAS-CoV procedente de la extracción catalogada bajo el n° 031589, tal como se ha definido anteriormente, correspondiendo dicha secuencia a aquella situada entre el nucleótido en la posición 28933 a 29727 (SEC ID n° 40), en referencia a la secuencia Genbank n° de acceso AY274119.3, se termina por una serie de nucleótidos a.

2.6) ORF1a y ORF1b

La amplificación de la región 5' que contiene los ORF1a y ORF1b del genoma del SRAS-CoV procedente de la extracción n° 031589 se realizó practicando unas reacciones de RT-PCR seguidas de PCR anidadas según los mismos principios que los anteriormente descritos para los otros ORF. Los fragmentos amplificados están superpuestos sobre varias decenas de bases, permitiendo así la reconstrucción informática de la secuencia completa de esta parte del genoma. De media, los fragmentos amplificados son de dos kilobases.

Se amplificaron así 14 fragmentos superpuestos denominados L0 a L12 con la ayuda de los cebadores siguientes:

Tabla II: Cebadores utilizados para la amplificación de la región 5'(ORF1a y ORF1b)

REGIÓN AMPLIFICADA Y SECUENCIADA (no tiene en cuenta los cebadores)	Cebador sentido RT-PCR	Cebador antisentido RT-PCR	Cebador sentido PCR anidada	Cebador antisentido PCR anidada
L0 50-480	S/L0/F1/+30	S/L0/R1/-481		
L1 231-2240	S/L1/F1/+147	S/L1/R1/-2336	S/L1/F2/+211	S/L1/R2/-2241
L2 2156-4167	S/L2/F1/+2033	S/L2/R1/-4192	S/L2/F2/+2136	S/L2/R2/-4168
L3 3913-5324	S/L3bis/F1/+3850	S/L3bis/R1/-5365	S/L3bis/F2/+3892	S/L3bis/R2/-5325
L4b 4952-6023	S/L4b/F1/+4878	S/L4b/R1/-6061	S/L4b/F2/+4932	S/L4b/R2/-6024
L4 5325-7318	S/L4/F1/+5272	S/L4/R1/-7392	S/L4/F2/+5305	S/L4/R2/-7323
L5 7296-9156	S/L5/F1/+7111	S/L5/R1/-9253	S/L5/F2/+7275	S/L5/R2/-9157
L6 9053-11066	S/L6/F1/+8975	S/L6/R1/-11151	S/L6/F2/+9032	S/L6/R2/-11067
L7 10928-12962	S/L7/F1/+10883	S/L7/R1/-13050	S/L7/F2/+10928	S/L7/R2/-12963
L8 12835-14834	S/L8/F1/+12690	S/L8/R1/-14857	S/L8/F2/+12815	S/L8/R2/-14835
L9 14765-16624	S/L9/F1/+14688	S/L9/R1/-16678	S/L9/F2/+14745	S/L9/R2/-16625
L10 16534-18570	S/L10/F1/+16451	S/L10/R1/-18594	S/L10/F2/+16514	S/L10/R2/-18571
L11 18521-20582	S/L11/F1/+18441	S/L11/R1/-20612	S/L11/F2/+18500	S/L11/R2/-20583
L12 20338-22205	S/L12/F1/+20279	S/L12/R1/-22229	S/L12/F2/+20319	S/L12/R2/-22206

Todos los fragmentos se purificaron en las condiciones siguientes, excepto el fragmento L0 que se amplificó como se describe anteriormente para el ORF-M:

- **RT-PCR:** 30 min. a 42°C, 15 min. a 55°C, 2 min. a 94°C, después el ADNc obtenido se amplifica en las condiciones siguientes: 40 ciclos que comprenden: una etapa de desnaturalización a 94°C durante 15 s, una etapa de hibridación a 58°C durante 30 s y después una etapa de elongación a 68°C durante 1 min. 30 s, con 5 s de elongación suplementaria en cada ciclo, después una etapa final de elongación a 68°C durante 7 min.

- **PCR anidada:** una etapa inicial de desnaturalización a 94°C durante 2 min. seguida de 35 ciclos que comprenden: una etapa de desnaturalización a 94°C durante 15 s, una etapa de hibridación a 60°C durante 30 s y después una etapa de elongación a 72°C durante 1 min. 30 s, con 5 s de elongación suplementaria en cada ciclo, después una etapa final de elongación a 72°C durante 7 min.

Los productos de amplificaciones se secuenciaron con la ayuda de los cebadores definidos en la Tabla III siguiente:

Tabla III: Cebadores utilizados para la secuenciación de la región 5' (ORF1a y ORF1b)

Nombres	Secuencias (SEC ID nº 76 a 139)
S/L3/+4932	5'-CCACACACAGCTTGTGGATA-3'
S/L4/+6401	5'-CCGAAGTTGTAGGCAATGTC-3'
S/L4/+6964	5'-TTTGGTGCTCCTTCTTATTG-3'
S/L4/-6817	5'-CCGGCATCCAAACATAATTT- 3'
S/L5/-7633	5'-TGGTCAGTAGGGTTGATTGG-3'
S/L5/-8127	5'-CATCCTTTGTGTCAACATCG-3'
S/L5/-8633	5'-GTCACGAGTGACACCATCCT-3'
S/L5/+7839	5'-ATGCGACGAGTCTGCTTCTA-3'
S/L5/+8785	5'-TTCATAGTGCCTGGCTTACC-3'
S/L5/+8255	5'-ATCTTGCGCATGTATTGAC-3'
S/L6/-9422	5'-TGCATTAGCAGCAACAACAT-3'

Nombres	Secuencias (SEC ID nº 76 a 139)
S/L6/-/9966	5'-TCTGCAGAACAGCAGAAGTG-3'
S/L6/-/10542	5'-CCTGTGCAGTTTGTCTGTCA-3'
S/L6/+10677	5'-CCTTGTGGCAATGAAGTACA-3'
S/L6/+10106	5'-ATGTCATTTGCACAGCAGAA-3'
S/L6/+9571	5'-CTTCAATGGTTTGCCATGTT-3'
S/L7/-/11271	5'-TGCGAGCTGTCATGAGAATA-3'
S/L7/-/11801	5'-AACCGAGAGCAGTACCACAG-3'
S/L7/-/12383	5'-TTTGGCTGCTGTAGTCAATG-3'
S/L7/+12640	5'-CTACGACAGATGTCCTGTGC-3'
S/L7/+12088	5'-GAGCAGGCTGTAGCTAATGG-3'
S/L7/+111551	5'-TTAGGCTATTGTTGCTGCTG-3'
S/L8/-13160	5'-CAGACAACATGAAGCACCAC-3'
S/L8/-/13704	5'-CGCTGACGTGATATATGTGG-3'
S/L8/-14284	5'-TGCACAATGAAGGATACACC-3'
S/L8/+14453	5'-ACATAGCTCGCGTCTCAGTT-3'
S/L8/+13968	5'-GGCATTGTAGGCGTACTGAC-3'
S/L8/+13401	5'-GTTTGCGGTGTAAGTGCAG-3'
S/L9/-15098	5'-TAGTGGCGGCTATTGACTTC-3'
S/L9/-15677	5'-CTAAACCTTGAGCCGCATAG-3'
S/L9/-16247	5'-CATGGTCATAGCAGCACTTG-3'
S/L9/+16323	5'-CCAGGTTGTGATGTCAGTGAT-3'
S/L9/+15858	5'-CCTTACCCAGATCCATCAAG-3'
S/L9/+15288	5'-CGCAAACATAACACTTGCTG-3'
S/L10/-16914	5'-AGTGTGGGTACAAGCCAGT-3'
S/L10/-17466	5'-GTTCCAAGGAACATGTCTGG-3'
S/L10/-18022	5'-AGGTGCCTGTGTAGGATGAA-3'
S/L10/+18245	5'-GGGCTGTCATGCAACTAGAG-3'
S/L10/+17663	5'-TCTTACACGCAATCCTGCTT-3'
S/L10/+17061	5'-TACCCATCTGCTCGCATAGT-3'
S/L11/-/18877	5'-GCAAGCAGAATTAACCCTCA-3'
S/L11/-19396	5'-AGCACCACCTAAATTGCATC-3'
S/L11/-20002	5'-TGGTCCCTTTGAAGGTGTTA-3'
S/L11/+20245	5'-TCGAACACATCGTTTATGGA-3'
S/L11/+19611	5'-GAAGCACCTGTTTCCATCAT-3'
S/L11/+19021	5'-ACGATGCTCAGCCATGTAGT-3'
SRAS/L1/F3/+800	5'-GAGGTGCAGTCACTCGCTAT-3'
SRAS/L1/F4/+1391	5'-CAGAGATTGGACCTGAGCAT-3'
SRAS/L1/F5/+1925	5'-CAGCAAACCACTCAATTCCT-3'
SRAS/L1/R3/-1674	5'-AAATGATGGCAACCTCTTCA-3'
SRAS/L1/R4/-1107	5'-CACGTGGTTGAATGACTTTG-3'
SRAS/L1/R5/-520	5'-ATTTCTGCAACCAGCTCAAC-3'
SRAS/L2/F3/+2664	5'-CGCATTGTCTCCTGGTTTAC-3'
SRAS/L2/F4/+3232	5'-GAGATTGAGCCAGAACCAGA-3'
SRAS/L2/F5/+3746	5'-ATGAGCAGGTTGTCATGGAT-3'
SRAS/L2/R3/-3579	5'-CTGCCTTAAGAAGCTGGATG-3'
SRAS/L2/R4/-2991	5'-TTTCTTACCAGCATCATCA-3'
SRAS/L2/R5/-2529	5'-CACCGTTCTTGAGAACAACC-3'
SRAS/L3/F3/+4708	5'-TCTTTGGCTGGCTCTTACAG-3'
SRAS/L3/F4/+5305	5'-GCTGGTGATGCTGCTAACTT-3'
SRAS/L3/F5/+5822	5'-CCATCAAGCCTGTGTCGTAT-3'
SRAS/L3/R3/-5610	5'-CAGGTGGTGCAGACATCATA-3'
SRAS/L3/R4/-4988	5'-AACATCAGCACCATCCAAGT-3'
SRAS/L3/R5/-4437	5'-ATCGGACACCATAGTCAACG-3'

Las secuencias de los fragmentos L0 a L12 de la cepa de SRAS-CoV procedente de la extracción catalogada bajo el nº 031589, corresponden respectivamente a las secuencias SEC ID nº 41 a SEC ID nº 54 en el listado de secuencias adjunto en anexo. Entre estas secuencias, sólo las que corresponden a los fragmentos L5 comprenden una diferencia nucleotídica con referencia a la secuencia correspondiente del aislado AY278741-Urbani. Esta mutación t/c en la posición 7919 desemboca una modificación de la secuencia en aminoácidos de la proteína correspondiente, codificada por el ORF 1a: en la posición 2552, una valina (codón gtt; AY278741) se cambia a alanina (codón gct) en la cepa de SRAS-CoV 031589. Sin embargo, no se identifica ninguna mutación con referencia a la secuencia correspondiente del aislado AY274119.3-Urbani. Los otros fragmentos no presentan diferencias con respecto a las secuencias correspondientes de los aislados Tor2 y Urbani.

Ejemplo 2: Producción y purificación de proteínas N y S recombinantes de la cepa de SRAS-CoV procedente de la extracción catalogada bajo el número 031589

La proteína entera N y dos fragmentos polipeptídicos de la proteína S de la cepa de SRAS-CoV procedente de la extracción catalogada bajo el número 031589 se produjeron en *E. coli*, en forma de proteínas de fusión que comprende una etiqueta polihistidina N- o C-terminal. En los dos polipéptidos S, las secuencias hidrófobas N y C-terminales de la proteína S (péptido señal: posiciones 1 a 13 y hélice transmembranaria: posiciones 1196 a 1218) se eliminaron mientras que se preservaron la hélice β (posiciones 565 a 687) y las dos unidades de tipo "hélices enrolladas" (posiciones 895 a 980 y 1155 a 1186) de la proteína S. Estos dos polipéptidos están constituidos por: un fragmento largo (S_L) que corresponde a las posiciones 14 a 1193 de la secuencia en aminoácidos de la proteína S y un fragmento corto (S_C) que corresponde a las posiciones 475 a 1193 de la secuencia en aminoácidos de la proteína S.

1) Clonación de los ADNc N, S_L y S_C en el vector de expresión pIVEX2.3 y pIVEX2.4

Los ADNc que corresponden a la proteína N y a los fragmentos S_L y S_C se amplificaron por PCR en condiciones estándares, con la ayuda del ADN polimerasa Platinum Pfx[®] (INVITROGEN). Los plásmidos SRAS-N y SRAS-S se usaron como plantillas y los oligonucleótidos siguientes como cebadores:

5'-CCCATATGTCTGATAATGGACCCCAATCAAAC-3' (N sentido, SEC ID n° 55)

5'-CCCCCGGGTGCCTGAGTTGAATCAGCAGAAGC-3' (N antisentido, SEC ID n° 56)

5'-CCCATATGAGTGACCTTGACCGGTGCACCAC-3' (S_C sentido, SEC ID n° 57)

5'-CCCATATGAAACCTTGACCCCCACCTGCTC-3' (S_L sentido, SEC ID n° 58)

5'-CCCCCGGGTTTAATATATTGCTCATATTTTCCC-3' (S_C y S_L antisentido, SEC ID n° 59).

Los cebadores sentido introducen un sitio *NdeI* (subrayado) mientras que los cebadores antisentido introducen un sitio *XmaI* o *SmaI* (subrayado). Los 3 productos de amplificación se purificaron sobre columna (kit *QIAquick PCR Purification*, QIAGEN) y se clonaron en un vector apropiado. El ADN plasmídico purificado de las 3 construcciones (kit *QIAfilter Midi Plasmid*, QIAGEN) se verificó mediante secuenciación y se digirió mediante las enzimas *NdeI* y *XmaI*. Los 3 fragmentos que corresponden a los ADNc N, S_L y S_C se purificaron sobre gel de agarosa y después se insertaron en los plásmidos pIVEX2.3MCS (etiqueta polihistidina C-terminal) y pIVEX2.4d (etiqueta polihistidina N-terminal) previamente digeridos por las mismas enzimas. Después de la verificación de las construcciones, los 6 vectores de expresión así obtenidos (pIV2.3N, pIV2.3 S_C , pIV2.3 S_L , pIV2.4N, pIV2.4 S_C también denominado pIV2.4 S_i , pIV2.4 S_L) se usaron después, por una parte para ensayar la expresión de las proteínas *in vitro*, y por otra parte para transformar la cepa bacteriana BL21(DE3)pDIA17 (NOVAGEN). Estas construcciones codifican unas proteínas cuya masa molecular esperada es la siguiente: pIV2.3N (47174 Da), pIV2.3 S_C (82897 Da), pIV2.3 S_L (132056 Da), pIV2.4N (48996 Da), pIV2.4 S_C (81076 Da) y pIV2.4 S_L (133877 Da). Unas bacterias transformadas por pIV2.3N se depositaron en la CNCM el 23 de octubre de 2003, bajo el número I-3117, y unas bacterias transformadas por pIV2.4 S_i se depositaron a la CNCM el 23 de octubre de 2003, bajo el número I-3118.

2) Análisis de la expresión de las proteínas recombinantes *in vitro* e *in vivo*

Se ha ensayado la expresión de proteínas recombinantes a partir de los 6 vectores recombinantes, en una primera etapa, en un sistema *in vitro* (RTS100, Roche). Las proteínas producidas *in vitro*, después de una incubación de los vectores recombinantes pIVEX, 4h a 30°C, en el sistema RTS100, se analizaron mediante transferencia western con la ayuda de un anticuerpo anti-(his)₆ acoplado a la peroxidasa. El resultado de la expresión *in vitro* (figura 1) muestra que sólo la proteína N está expresada en cantidades importantes, sea cual sea la posición, N- o C-terminal, de la etiqueta polihistidina. En una segunda etapa, la expresión de las proteínas N y S se ensayó *in vivo* a 30°C en un medio LB, en presencia o ausencia de inductor (IPTG 1 mM). La proteína N está muy bien producida en este sistema bacteriano (figura 2) y se encuentra principalmente en una fracción soluble después de la lisis de las bacterias. Sin embargo, la versión larga de S (S_L) está muy poco producida y completamente insoluble (figura 3). La versión corta (S_C) presenta también una muy baja solubilidad, pero un índice de expresión mucho más elevado que el de la versión larga. Por otra parte, la construcción S_C fusionada con una etiqueta polihistidina en posición C-terminal presenta un tamaño más bajo de lo que se espera. Un experimento de inmunodetección con un anticuerpo anti-polihistidina ha mostrado que esta construcción era incompleta. En conclusión, las dos construcciones pIV2.3N y pIV2.4 S_i que expresan respectivamente la proteína N entera fusionada con la etiqueta polihistidina en C-terminal y la proteína S corta fusionada con la etiqueta polihistidina en N-terminal, se seleccionaron para producir las dos proteínas en gran cantidad a fin de purificarlas. Los plásmidos pIV2.3N y pIV2.4 S_i se han depositado respectivamente bajo el n° I-3117 y I-3118 en la CNCM, 25 rue du Docteur Roux, 75724 PARIS 15, el 23 de octubre de 2003.

3) Análisis de la actividad antigénica de las proteínas recombinantes

Se ha ensayado la actividad antigénica de las proteínas N, S_L y S_C mediante transferencia western, con la ayuda de dos muestras de suero, que provienen de un mismo paciente infectado por el SRAS-CoV, extraídas 8 días (M12) y 29 días (M13) después del principio de los síntomas del SRAS. El protocolo experimental es como se describe en el ejemplo 3. Los resultados ilustrados por la figura 4 muestran (i) la seroconversión del paciente, y (ii) que la proteína N posee una mayor reactividad antigénica que la proteína S corta.

4) Purificación de la proteína N a partir de pIV2.3N

Se han realizado varios experimentos de purificación de la proteína N, producida a partir del vector pIV2.3N, según el protocolo siguiente. Las bacterias BL21(DE3)pDIA17, transformadas por el vector de expresión pIV2.3N, se cultivaron a 30°C en 1 litro de medio de cultivo que contiene 0,1 mg/ml de ampicilina, e inducidas por 1 mM de IPTG cuando se alcanza la densidad celular, equivalente a A₆₀₀ = 0,8 (aproximadamente 3 horas). Después de 2 horas de cultivo en presencia de inductor, las células se recuperaron mediante centrifugación (10 min. a 5000 rpm), se resuspendieron en el tampón de lisis (50 mM NaH₂PO₄, NaCl 0,3 M, 20 mM imidazol, pH 8 que contiene la mezcla de inhibidores de proteasas *Complete*[®], Roche), y se lisaron mediante la prensa de French (12000 psi). Después de la centrifugación del lisado bacteriano (15 min. a 12000 rpm), el sobrenadante (50 ml) se depositó con un caudal de 1ml/min. sobre una columna (15 ml) de quelación metálica (Ni-NTA superflow, Qiagen), equilibrada por el tampón de lisis. Después del lavado de la columna por 200 ml de tampón de lisis, la proteína N se eluyó mediante un gradiente de imidazol (20 → 250 mM) en 10 volúmenes de columna. Las fracciones que contienen la proteína N se reunieron y se analizaron mediante electroforesis en gel de poliacrilamida en condiciones desnaturalizantes y después coloración con azul de Coomassie. Los resultados ilustrados por la figura 5 muestran que el protocolo empleado permite purificar la proteína N con una homogeneidad muy satisfactoria (95%) y un rendimiento medio de 15 mg de proteína por litro de cultivo.

5) Purificación de la proteína S_C a partir de pIV2.4S_C (pIV2.4S₁)

El protocolo seguido para purificar la proteína S corta es muy diferente de aquel descrito anteriormente, ya que la proteína está fuertemente agregada en el sistema bacteriano (cuerpo de inclusión). Las bacterias BL21(DE3)pDIA17, transformadas por el vector de expresión pIV2.4S₁ se cultivaron a 30°C en 1 litro de medio de cultivo que contiene 0,1 mg/ml de ampicilina, e inducidas mediante 1 mM de IPTG cuando se alcanza la densidad celular, equivalente a A₆₀₀ = 0,8, (aproximadamente 3 horas). Después de 2 horas de cultivo en presencia de inductor, las células se recuperaron mediante centrifugación (10 min. a 5000 rpm), se resuspendieron en el tampón de lisis (0,1 M Tris-HCl, EDTA 1 mM, pH 7,5), y se lisaron mediante la prensa de French (1200 psi). Después de la centrifugación del lisado bacteriano (15 min. a 12000 rpm), el residuo se resuspendió en 25 ml de tampón de lisis que contiene 2% de Triton X100 y 10 mM β -mercaptoetanol, y después se centrifugó durante 20 min. a 12000 rpm. El residuo se resuspendió en un tampón Tris-HCl 10 mM que contiene 7 M de urea, y se agitó suavemente durante 30 min. a temperatura ambiente. Este último lavado de los cuerpos de inclusión con 7 M de urea es necesario para eliminar la mayoría de las proteínas membranas de *E. coli* que cosedimentan con la proteína S_C agregada. Después de una última centrifugación durante 20 min. a 12000 rpm, el residuo final se resuspende en el tampón Tris-HCl 10 mM. El análisis electroforético de esta preparación (figura 6) muestra que la proteína S corta puede ser purificada con una homogeneidad satisfactoria (aproximadamente el 90%) a partir de los cuerpos de inclusión (extracto insoluble).

Ejemplo 3: Inmunodominancia de la proteína N

La reactividad de los anticuerpos presentes en el suero de los pacientes que padecen neumopatía atípica causada por el coronavirus asociado al SRAS (SRAS-CoV), frente a diferentes proteínas de este virus, se analizó mediante transferencia western en las condiciones descritas a continuación.

1) Material

a) lisado de células infectadas por el SRAS-CoV

Células Vero E6 (2x10⁶) se infectaron por el SRAS-CoV (aislado catalogado bajo el número FFM/MA104) a una multiplicidad de infección (M.O.I.) de 10⁻¹ o 10⁻² y después se incubaron en un medio DMEM que contiene el 2% de SVF, a 35°C en una atmósfera que contiene 5% de CO₂. 48 horas más tarde, la alfombra celular se lavó con PBS y después se lisó con 500 μ l de tampón de depósito preparado según Laemmli, y que contiene β -mercaptoetanol. Las muestras se hirvieron después durante 10 minutos y después se sonificaron 3 veces durante 20 segundos.

b) anticuerpos**b₁) suero de paciente que padece neumopatía atípica**

El suero referenciado en el Centre National de Référence des virus influenzae (Région-Nord) bajo el n° 20033168 es el de un paciente francés que padece neumopatía atípica causada por el SRAS-CoV extraído 38 días después del principio de los síntomas; el diagnóstico de infección por el SRAS-CoV se realizó mediante par RT-PCR anidada y PCR cuantitativa.

b₂) sueros policlonales de conejo monoespecíficos dirigidos contra la proteína N o la proteína S

Los sueros son los producidos a partir de las proteínas recombinantes N y S_C (ejemplo 2), según el protocolo de inmunización descrito en el ejemplo 4; se trata del suero de conejo P13097 (suero anti-N) y del suero de conejo P11135 (suero anti-S).

2) Método

Se separaron 20 µl de lisado de células infectadas por el SRAS-CoV con M.O.I. de 10^{-1} y 10^{-2} y, a título de control, 20 µl de un lisado de células no infectadas (mock) sobre un gel SDS al 10% de poliacrilamida y después se transfirieron sobre una membrana de nitrocelulosa. Después del bloqueo en una solución de PBS/leche 5%/Tween 0,1% y lavado en PBS/ Tween 0,1%, esta membrana se hibridó durante una noche a 4°C con: (i) el inmunosuero n° 20033168 diluido al 1/300, 1/1000 y 1/3000 en el tampón PBS/BSA 1%/Tween 0,1%, (ii) el suero de conejo P13097 (suero anti-N) diluido al 1/50000 en el mismo tampón y (iii) el suero de conejo P11135 (suero anti-S) diluido al 1/10000 en el mismo tampón. Después del lavado en PBS/Tween, se realizó una hibridación secundaria con la ayuda de anticuerpos policlonales de carnero dirigidos contra las cadenas pesadas y ligeras de las inmunoglobulinas G humanas y acoplados a la peroxidasa (NA933V, Amersham), o bien de anticuerpos policlonales de asno dirigidos contra las cadenas pesadas y ligeras de las inmunoglobulinas G de conejo y acoplados a la peroxidasa (NA934V, Amersham). Los anticuerpos fijados se revelaron con la ayuda del kit ECL+ (Amersham) y de películas de autorradiografía Hyperfilm MP (Amersham). Una escala de masa molecular (kDa) se muestra en la figura.

3) Resultados

La figura 7 muestra que tres polipéptidos de masa molécula aparente 35, 55 y 200 kDa son detectados específicamente en los extractos de células infectadas por el SRAS-CoV.

A fin de identificar estos polipéptidos, se realizaron otras dos inmunohuellas (figura 8) sobre las mismas muestras y en las mismas condiciones con dos anticuerpos policlonales de conejo específicos de la nucleoproteína N (conejo P13097, figura 8A) y de la proteína de espícula S (conejo P11135, figura 8B). Este experimento muestra que el polipéptido de 200 kDa corresponde a la glicoproteína de espícula S del SRAS-CoV, que el polipéptido de 55 kDa corresponde a la nucleoproteína N mientras que el polipéptido de 35 kDa representa posiblemente una forma truncada o degradada de la N.

Los datos presentados en la figura 7 muestran por lo tanto que el suero 20033168 reacciona fuertemente con la N, y mucho más débilmente con la S del SRAS-CoV, ya que los polipéptidos de 35 y 55 kDa se revelan en forma de bandas intensas para diluciones de 1/300, 1/1000 y 1/3000 del inmunosuero, mientras que el polipéptido de 200 kDa es débilmente revelado para una dilución de 1/300. Se puede anotar también que ningún otro polipéptido del SRAS-CoV es detectado para diluciones superiores al 1/300 del suero 20033168.

Este experimento indica que la respuesta en anticuerpo específico de la N del SRAS-CoV domina las respuestas en anticuerpos específicos de los otros polipéptidos del SRAS-CoV y en particular la respuesta en anticuerpos dirigida contra la glicoproteína S. Indica una inmunodominancia de la nucleoproteína N durante infecciones humanas por el SRAS-CoV.

Ejemplo 4: Preparación de anticuerpos policlonales monoespecíficos dirigidos contra las proteínas N y S del coronavirus asociado al SRAS (SRAS-CoV)**1) Material y método**

Se inmunizaron tres conejos (P13097, P13081, P13031) con el polipéptido recombinante purificado que corresponde a la totalidad de la nucleoproteína (N), preparado según el protocolo descrito en el ejemplo 2. Después de una primera inyección de 0,35 mg por conejo de proteína emulsionada en adyuvante completo de Freund (vía intradérmica), los animales recibieron 3 inyecciones de recuerdo con 3 y después con 4 semanas de intervalo, de 0,35 mg de proteína recombinante emulsionada en adyuvante completo de Freund.

Se inmunizaron tres conejos (P11135, P13042, P14001) con el polipéptido recombinante que corresponde al fragmento corto de la proteína S (S_C), producida como se describe en el ejemplo 2. Como este polipéptido se

encuentra principalmente en forma de cuerpo de inclusión en el citoplasma bacteriano, los animales han recibido 4 inyecciones intradérmicas con 3-4 semanas de intervalo de una preparación de cuerpo de inclusión que corresponde a 0,5 mg de proteína recombinante emulsionada en adyuvante incompleto de Freund. Las 3 primeras inyecciones se realizaron con una preparación de cuerpos de inclusión, preparados según el protocolo descrito en el ejemplo 2, mientras que la cuarta inyección se realizó con una preparación de cuerpos de inclusión, preparados según el protocolo descrito en el ejemplo 2, y después purificados sobre gradiente de sacarosa y lavados en 2% de Tritón X100.

Para cada conejo, se ha preparado un suero preinmune (p.i.) antes de la primera inmunización y un inmunosuero (I.S.) 5 semanas después de la cuarta inmunización.

En una primera etapa, se analizó la reactividad de los sueros mediante ensayo ELISA frente a preparaciones de proteínas recombinantes parecidas a las utilizadas para las inmunizaciones; los ensayos ELISA se realizaron según el protocolo y con los agentes reactivos descritos en el ejemplo 6.

En una segunda etapa, se analizó la reactividad de los sueros realizando una inmunohuella (Western-blot) de un lisado de células infectadas por el SRAS-CoV, siguiendo el protocolo tal como se ha descrito en el ejemplo 3.

2) Resultados

Los ensayos ELISA (figura 9) demuestran que las preparaciones de proteína N recombinante y de cuerpos de inclusión del fragmento corto de la proteína S (S_C) son inmunógenos en el animal y que el título de los sueros inmunes es elevado (más del 1/25000).

La inmunohuella (figura 8) muestra que el suero inmune del conejo P13097 reconoce dos polipéptidos presentes en los lisados de células infectadas por el SRAS-CoV: un polipéptido cuya masa molecular aparente (50-55 kDa según los experimentos) es compatible con la de la nucleoproteína N (422 residuos, masa molecular predicha de 46 kDa) y un polipéptido de 35 kDa, que representa probablemente una forma truncada o degradada de la N.

Este experimento muestra también que el suero de conejo P11135 reconoce principalmente un polipéptido cuya masa molecular aparente (180-220 kDa según los experimentos) es compatible con una forma glicosilada de la S (1255 residuos, cadena polipeptídica no glicosilada de 139 kDa), así como unos polipéptidos más ligeros, que representan probablemente unas formas truncadas y/o glicosiladas de la S.

En conclusión, el conjunto de estos experimentos demuestra que unos polipéptidos recombinantes expresados en *E. coli* y que corresponden a las proteínas N y S del SRAS-CoV permiten inducir en el animal unos anticuerpos policlonales capaces de reconocer las formas nativas de estas proteínas.

Ejemplo 5: Preparación de anticuerpos policlonales mono-específicos dirigidos contra las proteínas M y E del coronavirus asociado al SRAS (SRAS-CoV)

1) Análisis de la estructura de las proteínas M y E

a) Proteína E

La estructura de la proteína E del SRAS-CoV (76 aminoácidos) se analizó *in silico*, con la ayuda de diferentes programas como signalP v1.1, NetNGlyc 1.0, THMM 1.0 y 2.0 (Krogh y otros., 2001, J. Mol. Biol., 305(3):567-580) o también TOPPRED (von Heijne, 1992, J. Mol. Biol. 225, 487-494). El análisis muestra que este polipéptido no glicosilado es una proteína membranaria de tipo 1, que contiene una sola hélice transmembranaria (aa 12-34 según THMM), y cuya parte más grande del dominio hidrófilo (42 residuos) está localizada en el extremo C-terminal y posiblemente en el interior de la partícula viral (endodominio). Se puede anotar una inversión en la topología predicha por las versiones 1.0 (N-ter es externo) y 2.0 (N-ter es interno) del programa THMM, pero que otros algoritmos, en particular TOPPRED y THUMBUP (Zhou y Zhou, 2003, Protein Science 12:1547-1555) confirman una localización externa del extremo N-terminal de E.

b) Proteína M

Un análisis similar realizado sobre la proteína M del SRAS-CoV (221 aminoácidos) muestra que este polipéptido no posee péptido señal (según el programa signalP v1.1) sino tres dominios transmembranarios (residuos 15-37, 50-72, 77-99 según THMM2.0) y un gran dominio hidrófilo (aa 100-221) localizado en el interior de la partícula viral (endodominio). Esta posiblemente glicosilada sobre la asparagina en la posición 4 (según NetNGlyc 1.0).

Así, de acuerdo con los datos experimentales conocidos para los otros coronavirus, es remarcable que las dos proteínas M y E presentan unos endodominios que corresponden a la mayor parte de sus polipéptidos y unos ectodominios de tamaño muy pequeño.

- el ectodominio de E corresponde posiblemente a los residuos 1 a 11 ó 1 a 12 de la proteína: MYSFVSEETGT (L), SEC ID nº 70. En efecto, la probabilidad asociada a la localización transmembranaria del residuo 12 es intermedia (0,56 según THMM 2.0).

- el ectodominio de M corresponde posiblemente a los residuos 2 a 14 de la proteína: ADNGTITVEELKQ, SEC ID nº 69. En efecto, la metionina N-terminal de M está, muy probablemente, escindida del polipéptido maduro, ya que el residuo en la posición 2 es una alanina (Varshavsky, 1996, 93:12142-12149).

Por otra parte, el análisis de hidrofobicidad (Kyte & Doolittle, Hopp & Woods) de la proteína E demuestra que el extremo C-terminal del endodominio de E es hidrófilo y por lo tanto probablemente expuesto a la superficie de este dominio. Así, un péptido sintético que corresponde a este extremo es un buen candidato inmunógeno para inducir en el animal unos anticuerpos dirigidos contra el endodominio de E. En consecuencia, se ha sintetizado un péptido que corresponde a los 24 residuos C-terminales de E.

2) Preparación de anticuerpos dirigidos contra el ectodominio de las proteínas M y E y el endodominio de la proteína E

Los péptidos M2-14 (ADNGTITVEELKQ, SEC ID nº 69), E1-12 (MYSFVSEETGTL, SEC ID nº 70) y E53-76 (KPTVYVYSRV KNLNSSEGV DLLV, SEC ID nº 71) se sintetizaron mediante Neosystem. Se han acoplado a la KLH (*Keyhole Limpet Hemocyanin*) con la ayuda del MBS (m-maleimido-benzoil-N-hidroxisuccinimida éster) a través de una cisteína añadida durante la síntesis, o bien en N-terminal del péptido (caso de E53-76), o bien en C-terminal (caso de M2-14 y E1-12).

Se inmunizaron dos conejos con cada uno de los conjugados, siguiendo el protocolo de inmunización siguiente: después de una primera inyección de 0,5 mg de péptido acoplado a la KLH y emulsionado en adyuvante completo de Freund (vía intradérmica), los animales reciben 2 a 4 inyecciones de recuerdo con 3 ó 4 semanas de intervalo de 0,25 mg de péptido acoplado a la KLH y emulsionado en adyuvante incompleto de Freund.

Para cada conejo, se ha preparado un suero preinmune (p.i.) antes de la primera inmunización, y se prepara un inmunosuero (I.S.) 3 a 5 semanas después de la inyección de recuerdo.

La reactividad de los sueros se analizó mediante transferencia western con la ayuda de extractos de células infectadas por el SRAS-CoV (figura 43B), o con la ayuda de extractos de células infectadas por un virus recombinante de la vacuna que expresa la proteína E (VV-TG-E, figura 43A) o M (VV-TN-M, figura 43C) del aislado 031589 del SRAS-CoV.

Los inmunosueros de los conejos 22234 y 22240, inmunizados por el conjugado KLH-E53-76, reconocen un polipéptido de aproximadamente 9 a 10kD, que está presente en los extractos de células infectadas por el SRAS-CoV pero ausente en los extractos de células no infectadas (figura 43B). La masa aparente de este polipéptido es compatible con la masa predicha de la proteína E, que es de 8,4 kD. De manera similar, el inmunosuero del conejo 20047, inmunizado por el conjugado KLH-E1-12, reconoce un polipéptido presente en los extractos de células infectadas por el virus VV-TG-E, cuya masa molar aparente es compatible con la de la proteína E (figura 43A).

El inmunosuero de los conejos 20013 y 20080, inmunizados por el conjugado KLH-M2-14, reconoce un polipéptido presente en los extractos de células infectadas por el virus VV-TN-M (figura 43C), cuya masa molar aparente (18 kD aproximadamente) es compatible con la de la glicoproteína M, que es de 25,1 kD y presenta un punto isoelectrico elevado (9,1 para el polipéptido desnudo).

Estos resultados demuestran que los péptidos E1-12 y E53-76 por un lado, y el péptido M2-14 por otro lado, permiten inducir en el animal unos anticuerpos policlonales capaces de reconocer las formas nativas de las proteínas E y M respectivamente del SRAS-CoV.

Ejemplo 6: Análisis de la reactividad en ELISA de la proteína N recombinante, frente a sueros de pacientes que padecen SRAS

1) Material

El antígeno utilizado para preparar las fases sólidas es la nucleoproteína N recombinante purificada preparada según el protocolo descrito en el ejemplo 2.

Los sueros a ensayar (Tabla IV) se seleccionaron en base a unos resultados de análisis de su reactividad por inmunofluorescencia (título IF-SRAS), frente a células infectadas por el SRAS-CoV.

Tabla IV: Sueros ensayados en ELISA

Referencia	Nº de suero	Tipo de suero	Fecha del suero***	Título IF-SRAS
3050	A	Control	na*	nt**
3048	B	Control	na	nt
033168	D	Paciente 1- SRAS	27/04/03 (D38)	320
033397	E	Paciente-1 SRAS	11/05/03 (D52)	320
032632	F	Paciente -2 SRAS	21/03/03 (D17)	2500
032791	G	Paciente -3 SRAS	04/04/03 (D3)	<40
033258	H	Paciente -3 SRAS	28/04/03 (D27)	160
*na: no-aplicable. **nt: no-ensayado. *** los datos indicados corresponden al número de días después del principio de los síntomas de SRAS.				

2) Método

La proteína N (100 µl) diluida a diferentes concentraciones en tampón carbonato 0,1 M, pH 9,6 (1, 2 ó 4 µg/ml) se distribuye en los pocillos de placas ELISA, después se incuban las placas durante una noche a temperatura del laboratorio. Las placas se lavan con tampón PBS-Tween, se saturan con tampón PBS-leche desnatada-sacarosa (5%). Se añaden los sueros a ensayar (100 µl) previamente diluidos (1/50, 1/100, 1/200, 1/400, 1/800, 1/1600 y 1/3200), después se incuban las placas durante 1h a 37° C. Después de 3 lavados, se añade el conjugado anti-IgG humano marcado con peroxidasa (referencia 209-035-098, JACKSON) diluido al 1/18000, y se incuban las placas durante 1h a 37°C. Después de 4 lavados, se añade el cromógeno (TMB) y el sustrato (H₂O₂), y se incuban las placas durante 30 min. a temperatura ambiente, protegido de la luz. La reacción se detiene después y se mide la absorbencia a 450 nm con la ayuda de un lector automático.

3) Resultados

Los ensayos ELISA (figura 10) demuestran que la preparación de proteína N recombinante está reconocida específicamente por los anticuerpos de sueros de pacientes que padecen SRAS extraídos en fase tardía de la infección (≥ 17 días después del principio de los síntomas), mientras que no está reconocida de manera significativa por los anticuerpos de un suero de paciente extraído en fase precoz de la infección (3 días después del principio de los síntomas), ni por sueros controles de sujetos que no padecen SRAS.

Ejemplo 7: Ensayos ELISA realizados para una detección muy específica y sensible de una infección por el coronavirus asociado al SRAS, a partir de sueros de pacientes

1) Ensayo ELISA IgG indirecto

a) Reactivos

Preparación de las placas

Las placas son sensibilizadas por una solución de proteína N a 2 µg/ml en un tampón PBS 10mM pH 7,2, rojo de fenol a 0,25 ml/l. Se depositan 100 µl de solución en los pocillos y se dejan incubar a temperatura ambiente durante una noche. La saturación se realiza mediante un prelavado en tampón PBS 10 mM/tween 0,1%, seguido de un lavado con una solución de saturación PBS, 25% leche/sacarosa.

Diluyente de sueros

Tampón TRIS 0,48 g/l, PBS 10 mM, EDTA 3,7 g/l, leche 15% v/v, pH 6,7

Diluyente conjugado

Tampón citrato (15 g/l), tween 0,5%, suero bovino 25%, NaCl 12%, leche desnatada 6% v/v PH 6,5

Conjugado

Conjugado anti-IgG humano 50X, comercializado por Bio-Rad: kit Platelia H. pylori ref. 72778

Otras Soluciones:

Solución de lavado R2, solución de revelado al TMB R8 diluyente, R9 cromógeno, R10 solución de detención: agentes reactivos comercializados por Bio-Rad (ej.: kit Platelia pylori, ref. 72778)

b) Modo de realización

Diluir los sueros al 1/200 en el diluyente de las muestras

Distribuir 100 µl/pocillo

Incubación durante 1h a 37°C

3 lavados en solución de LAVADO R2 10x previamente diluido 10 veces en agua desmineralizada (es decir, solución de lavado 1X)

Distribuir 100 µL de conjugado (conjugado 50x para diluir extemporáneamente en el diluyente conjugado proporcionado)

Incubación durante 1h a 37°C

4 lavados en solución de lavado 1X

Distribuir 200 µl/pocillo de solución de revelado (a diluir extemporáneamente, ej.: 1 ml de R9 en 10ml de R8)

Incubación durante 30 min. a temperatura ambiente en la oscuridad

Detener la reacción con 100 µl/pocillo de R10

LECTURA a 450/620nm

Los resultados pueden ser interpretados tomando un suero LIMITE que da una respuesta más allá de la cual los sueros ensayados se considerarán como positivos. Este suero se selecciona y se diluye a fin de dar una señal significativamente superior al ruido de fondo.

2) Ensayo ELISA DOBLE EPÍTOPO

a) Agentes reactivos

Preparación de las placas

Las placas son sensibilizadas mediante una solución de proteína N a 1 µg/ml en un tampón PBS 10 mM pH 7,2, rojo de fenol a 0,25 ml/l. Se depositan 100 µl de solución en los pocillos y se dejan incubar a temperatura ambiente durante una noche. La saturación se lleva a cabo mediante un prelavado en tampón PBS 10 mM/0,1% tween seguido de un lavado con una solución de saturación PBS 10 mM, leche 25% (V/V)

Diluyente de sueros y conjugado

Tampón TRIS salino 50 mM pH 8, leche 2%

Conjugado

Se trata de la proteína N recombinante purificada, acoplada a la peroxidasa según el protocolo de Nakane (Nakane P.K. y Kawaoi A; (1974): Peroxydase-labeled antibody, a new method of conjugation. The Journal of Histochemistry and Cytochemistry Vol. 22, N) 23, p. 1084-1091), en unas relaciones molares respectivas 1/2. Este conjugado ProtN POD se utiliza a una concentración de 2 µg/ml en un diluyente suero/conjugado.

Otras soluciones:

Solución de lavado R2, soluciones de revelado con TMB R8, diluyente, R9 cromógeno, R10 solución de detención: agentes reactivos comercializados por Bio-Rad (ej. kit platelia pylori ref. 72778).

b) Modo de realización

1ª etapa en placa de "predilución"

- Diluir cada suero al 1/5 en la placa de predilución (48 µl de diluyente + 12 µl de suero).
- Después de diluir el conjunto de los sueros, distribuir 60 µl de conjugado

- Llegado el caso, se deja incubar la mezcla suero + conjugado.

2ª etapa en placa de "reacción"

- Transferir 100 µl de mezcla/pocillo en la placa de reacción
- Incubación durante 1h 37°C
- 5 lavados en disolución de LAVADO R2 10x previamente diluida 10 veces con agua desmineralizada (→ solución de lavado 1x)
- Distribuir 200 µl/pocillo de solución de revelado (a diluir extemporáneamente, ej.: 1 ml de R9 en 10 ml de R8)
- Incubación durante 30 min. a temperatura ambiente protegido de la luz
- Detener la reacción con 100 µl/pocillo de R10
- LECTURA a 450/620nm

Como para el ensayo ELISA indirecto, los resultados pueden ser interpretados utilizando un suero de "valor límite". Cualquier suero que tiene una respuesta superior al suero de valor límite se considerará como positivo.

2) Resultados

Los sueros de pacientes clasificados como posible caso de SRAS del hospital francés de Hanoi, Vietnam o en relación con el hospital francés de Hanoi (JYK) se han analizado utilizando el ensayo IgG-N indirecto y el ensayo N doble epítipo.

Los resultados del ensayo IgG-N indirecto (figuras 14 y 15) y N doble epítipo (figuras 16 y 17) muestran una excelente correlación entre sí, así como con un ensayo ELISA indirecto que compara la reactividad de los sueros frente a un lisado de células VeroE6 no infectadas o infectadas por el SRAS-CoV (ELISA-lisado SRAS-CoV; véase la tabla V a continuación). Todos los sueros extraídos 12 días o más después del principio de los síntomas se encontraron positivos, incluyendo en pacientes para los cuales la infección por el virus del SRAS-CoV no había podido ser documentada mediante análisis de extracción respiratoria por RT-PCR, probablemente debido a una extracción demasiado tardía durante la infección ($\geq$ D12). En el caso del paciente TTH para el cual se realizó una extracción nasal a D7 fue encontrado negativo mediante RT-PCR, la calidad de la extracción podría ser la causa.

Ciertos sueros fueron encontrados negativos mientras que la presencia de SRAS-CoV se detectó mediante RT-PCR. Se trata en todos los casos de sueros precoces extraídos menos de 10 días después del principio de los síntomas (ej.: suero nº 032637). En el caso de un paciente PTTH (suero nº 032673), sólo se evocó una suspicacia de SRAS en el momento en el que las extracciones se realizaron.

En conclusión, los ensayos serológicos IgG-N indirecto y N-doble epítipo permiten documentar la infección por el SRAS-CoV en todos los pacientes para sueros extraídos 12 días o más después de la infección.

Tabla V: resultados de los ensayos ELISA

Nº Pvt	Paciente	Día	PCR-SRAS (1)	ELISA lisado SRAS-CoV (2)	IgG-N (2ª serie)	2Xepítipo (2ª serie)
033168	JYK	38	POS	+++	>5000	NT
033597	JYK	74	POS	NT	≈ 5000	NT
032552	VTT	8	NEG-D3 y D8 y D12	NEG	<200	<5
032544	CTP	16	NEG D16 y D20	++	>5000	>>20
032546	CJF	15	NEG D15 y D19	++	>5000	>>20
032548	PTL	17	NEG D17 y D21	++	>5000	>>20
032550	NTH	17	NEG-D17 y D21	++	>5000	>>20
032553	VTT	8	NEG-D3 y D8 y D12	NEG	<200	<5
032554	NTBV	4	POS	NEG	<200	<5

Nº Pvt	Paciente	Día	PCR-SRAS (1)	ELISA lisado SRAS-CoV (2)	IgG-N (2ª serie)	2Xepitopo (2ª serie)
032555	NTBV	4	POS	NEG	<200	
032564	NTP	15	POS	++	>5000	>>20
032629	NVH	4	POS	NEG	<200	<5
032631	BTTX	9	POS	NEG	<200	<5
032635	NHH	4	POS	NEG	<200	<5
032637	NHB	10	POS	NEG	<200	<5
032642	BTTX	9	POS	NEG	<200	<5
032643	LTDH	1	POS	NEG	<200	<5
032644	NTBV	4	POS	NEG	<200	<5
032646	TTH	12	NEG D7 y D 12 y D16	++	>5000	>>20
032647	DTH	17	NEG D17 y D21	++	>5000	>>20
032648	NNT	15	NEG D15 y D19	++	>5000	>>20
032649	PTH	17	NEG D17 y D21	++	>5000	>>20
032672	LVV	16	NEG D16 y D20	+	>5000	>>20
032673	PTTH	NA	NEG	NEG	<200	<5
032674	PNB	17	NEG D17 y D21	++	>5000	>>20
032682	VTH	12	NEG D12 y D16	++	>5000	>>20
032683	DTV	17	NEG D17 y D21	+	>1000	>>20

Notas:

(1): Los análisis por RT-PCR se realizaron por RT-PCR anidada BNI, LC Artus y LC-N sobre unos hisopos nasales o de la faringe; POS significa que al menos una extracción se encontró positiva en este paciente.

(2): La reactividad de los sueros en el ensayo ELISA que utiliza un lisado de células infectadas por el SRAS-CoV se clasificó en muy altamente reactivo (+++), altamente activo (++) , reactivo (+) y negativo en función del valor DO obtenido en las soluciones ensayadas.

Ejemplo 8: Detección del coronavirus asociado al SRAS (SRAS-CoV) por RT-PCR

1) Desarrollo de condiciones de RT-PCR en tiempo real con la ayuda de cebadores específicos del gen de la proteína de nucleocapside - ensayo "Light Cycler N"

a) concepción de los cebadores y de las sondas

La concepción de los cebadores y de las sondas se realizó a partir de la secuencia del genoma de la cepa de SRAS-CoV procedente de la extracción catalogada bajo el número 031589, con la ayuda del programa "Light Cycler Probe Design (Roche)". Así, se seleccionaron las dos series de cebadores y de sondas siguientes:

- **serie 1 (SEC ID nº 60, 61, 64, 65):**

- cebador sentido: N/+28507: 5'-GGC ATC GTA TGG GTT G-3' [28507-28522]

- cebador antisentido: N/-28774: 5'-CAG TTT CAC CAC CTC C-3' [28774-28759]

- sonda 1: 5'-GGC ACC CGC AAT CCT AAT AAC AAT GC-fluoresceína 3' [28561-28586]

- sonda 2: 5' Red705 -GCC ACC GTG CTA CAA CTT CCT-fosfato [28588-28608]

- **serie 2 (SEC ID nº 62, 63, 66, 67)**

- cebador sentido: N/+28375: 5'-GGC TAC TAC CGA AGA G-3' [28375-28390]

- cebador antisentido: N/-/28702: 5'-AAT TAC CGC GAC TAC G-3' [28702-28687]
- sonda 1: SRAS/N/FL: 5'-ATA CAC CCA AAG ACC ACA TTG GC-fluoresceína 3' [28541-28563]
- sonda 2: SRAS/N/LC705: 5' Red705 -CCC GCA ATC CTA ATA ACA ATG CTG C-fosfato 3' [28565-28589]

b) análisis de la eficacia de los dos pares de cebadores

A fin de ensayar la eficacia respectiva de los dos pares de cebadores, se realizó una amplificación por RT-PCR sobre un ARN sintético que corresponde a los nucleótidos 28054-29430 del genoma de la cepa de SRAS-CoV procedente de la extracción catalogada bajo el número 031589 y que contiene la secuencia del gen N.

De manera más precisa:

Este ARN sintético se preparó mediante transcripción *in vitro* con la ayuda del ARN polimerasa del fago T7, de una plantilla de ADN obtenida por linearización del plásmido SRAS-N con la enzima *Bam H1*. Después de la eliminación de la plantilla de ADN por digestión con la ayuda de DNasa 1, los ARN sintéticos se purifican mediante una extracción con fenol-cloroformo seguida de dos precipitaciones sucesivas en acetato de amonio e isopropanol. Se cuantifican entonces mediante la medición de la absorbencia a 260 nm y se controla su calidad mediante la relación de las absorbencias a 260 y 280 nm así como mediante una electroforesis en gel de agarosa. Así, la concentración de la preparación de ARN sintético utilizada para estos estudios es de 1,6 mg/ml, lo que corresponde a $2,1 \cdot 10^{15}$ copias/ml de ARN.

Cantidades decrecientes de ARN sintético se amplificaron por RT-PCR con la ayuda del kit "Superscript™ One-Step RT-PCR with Platinum® Taq" y los pares de cebadores n° 1 (N/+28507, N/-28774) (figura 1A) y n° 2 (N/+28375, N/-28702) (figura 1B), siguiendo las indicaciones del proveedor. Las condiciones de amplificación utilizadas son las siguientes: el ADNc se sintetizó mediante incubación durante 30 min. a 45 °C, 15 min. a 55°C y después 2 min. a 94 °C, después se amplificó mediante 5 ciclos que comprenden: una etapa de desnaturalización a 94°C durante 15 s, una etapa de hibridación a 45°C durante 30 s, y después una etapa de elongación a 72°C durante 30 s, seguidas de 35 ciclos que comprenden: una etapa de desnaturalización a 94°C durante 15 s, una etapa de hibridación a 55°C durante 30 s, y después una etapa de elongación a 72°C durante 30 s, con 2 s de elongación suplementaria en cada ciclo, y de una etapa final de elongación a 72°C durante 5 min. Los productos de amplificación obtenidos se mantuvieron después a 10°C.

Los resultados presentados en la figura 11 muestran que el par de cebadores n° 2 (N/+28375, N/-28702) permite detectar hasta 10 copias de ARN (banda de baja intensidad) o 10^2 copias (banda de buena intensidad) contra 10^4 copias para el par de cebadores n° 1 (N/+28507, N/-28774). Los amplicones son respectivamente de 268 pb (par 1) y de 328 pb (par 2).

c) desarrollo de la RT-PCR en tiempo real

Se desarrolló una RT-PCR en tiempo real con la ayuda del par de cebadores n°2 y del par de sonda constituido por SRAS/N/FL y SRAS/N/LC705 (figura 2).

La amplificación se realizó sobre un LightCycler™ (Roche) con la ayuda del kit "Light Cyler RNA Amplification Kit Hybridization Probes" (referencia 2 015 145, Roche) en las condiciones optimizadas siguientes. Una mezcla de reacción que contiene: H₂O (6,8 µl), MgCl₂ 25 mM (0,8 µl, 4 µM final de Mg²⁺), mezcla de reacción 5X (4 µl), sonda SRAS/N/FL 3 µM (0,5 µl, 0,075 µM final), sonda SRAS/N/LC705 3 µM (0,5 µl, 0,075 µM final), cebador N/+28375 10 µM (1 µl, 0,5 µM final), cebador N/-28702 10 µM (1 µl, 0,5 µM final), mezcla de enzima (0,4 µl) y muestra (ARN viral, 5 µl) se amplificó siguiendo el programa siguiente:

- <u>Transcripción inversa</u> :	50°C	10:00 min.	modo de análisis: ninguno
- <u>Desnaturalización</u> :	95°C	30 s x1	modo de análisis: ninguno
- <u>Amplificación</u> :	95°C	2 s	} modo de análisis: cuantificación*} x45 rampa térmica 2,0°C/s}
	50°C	15 s	
	72°C	13 s	
- <u>Enfriamiento</u> :	40°C	30 s x1	modo de análisis: ninguno

* la medición de fluorescencia se hace al final de la hibridación y a cada ciclo (en modo SINGLE).

Los resultados presentados en la figura 12 muestran que este RT-PCR en tiempo real es muy sensible ya que permite detectar 10^2 copias de ARN sintético en 100% de las 5 muestras analizadas (29/29 muestras en 8

experimentos) y hasta 10 copias de ARN en 100% de las 5 muestras analizadas (40/45 muestras en 8 experimentos). Muestra también que este RT-PCR permite detectar la presencia del genoma del SRAS-CoV en una muestra y cuantificar el número de genomas presentes. A título de ejemplo, el ARN viral de un stock de SRAS-CoV cultivado sobre células Vero E6 se extrajo con la ayuda del kit "Qiaamp viral RNA extraction" (Qiagen), diluido a $0,05 \cdot 10^{-4}$ y analizado por RT-PCR en tiempo real según el protocolo descrito anteriormente; el análisis presentado en la figura 12 muestra que este stock de virus contiene $6,5 \cdot 10^9$ genomas-equivalentes/ml (geq/ml), lo que es muy similar al valor de $1,0 \cdot 10^{10}$ geq/ml medido con la ayuda del kit "RealArt™ HPA-Coronavirus LC RT PCR Reagents" comercializado por Artus.

2) Desarrollo de condiciones de RT-PCR anidada que tiene como diana el gen del ARN polimerasa - ensayo "RT-PCR anidada CDC (Centers for Disease Control and Prevention) /IP"

a) Extracción del ARN viral

Muestra clínica: QIAamp viral RNA Mini Kit (QIAGEN) según las indicaciones del fabricante, o una técnica equivalente. El ARN se evalúa en un volumen de 60 µl.

b) RT-PCR anidada "SNE/SAR"

Primera etapa: RT-PCR acoplado «SNE»

Se utilizó el kit Invitrogen "Superscript™ One-Step RT-PCR with Platinum® Taq", pero se le puede sustituir el kit "Titan" de Roche Boehringer con resultados similares.

Oligonucleótidos:

- SNE-S1 5' GGT TGG GAT TAT CCA AAA TGT GA 3'
- SNE-AS 1 5'GCA TCA TCA GAA AGA ATC ATC ATG 3'
- > Tamaño esperado: **440 pb**

1. Preparar una mezcla:

H ₂ O	6,5 µl
Reacción mezcla 2X	12,5 µl
Óligo SNE-S1 50 µM	0,2 µl
Óligo SNE-AS1 50 µM	0,2 µl
ARNsin 40 U/µl	0,12 µl
RT/Platinum Taq mix	0,5 µl

2. A 20 µl de la mezcla, añadir 5 µl de ARN y proceder a la amplificación sobre termociclador (condiciones ABI 9600)

2.1	45°C	30 min.	
	55°C	15 min.	
	94°C	2 min.	
2.2.	94°C	15 s	}
	45°C	30 s	} x 5 ciclos
	72°C	30 s	}
2.3.	94°C	15 s	}
	55°C	30 s	} x 35 ciclos
	72°C	30 s + 2 s/ciclo	}
2.4.	72°C	5 min.	
2.5	10°C	∞	

Conservación a +4°C.

La ARNsin (N2511/N2515) de Promega se utilizó como inhibidor de RNasa.

Unos ARN sintéticos sirvieron de control positivo. A título de control, 10^3 , 10^2 y 10 copias de ARN sintético R_{SNE} se amplificaron en cada experimento.

Segunda etapa: PCR anidada "STAR"

Oligonucleótidos:

- SAR1-S 5' CCT CTC TTG TTC TTG CTC GCA 3'
- SAR1-AS 5' TAT AGT GAG CCG CCA CAC ATG 3'
- > Tamaño esperado: **121 pb**

1. Preparar una mezcla:

H ₂ O	35,8 µl
Tampón Taq 10X	5 µl
MgCl ₂ 25 mM	4 µl
Mezcla dNTPs 5 mM	2 µl
Oligo SAR1-S 50 µM	0,5 µl
Oligo SAR1-AS 50 µM	0,5 µl
Taq ADN pol 5 U/µl	0,25 µl

Se utilizó el *AmpliTaq DNA Pol* de Applied Biosystems (tampón 10X sin MgCl₂, ref. 27216601).

2. A 48 µl de la mezcla, añadir 2 µl del producto de la primera PCR y proceder a la amplificación (condiciones ABI 9600):

2.1.	94°C	2 min.	
2.2.	94°C	30 s	}
	45°C	45 s	} x 5 ciclos
	72°C	30 s	}
2.3.	94°C	30 s	}
	55°C	30 s	} x 35 ciclos
	72°C	30 s + 1 s/cycle	}
2.4.	72°C	5 min.	
2.5	10°C	∞	

3. Analizar 10 µl del producto de reacción sobre gel "low-melting" (tipo Seakem GTG) a 3% de agarosa. La sensibilidad del ensayo anidada está, rutinariamente, en las condiciones descritas, de 10 copias de ARN.

4. Los fragmentos puede después ser purificados sobre QIAquick PCR kit (QIAGEN) y secuenciados con los oligo SAR1-S y SAR1-AS.

3) Detección del ARN del SRAS-CoV por PCR a partir de extracciones respiratorias

a) Primer estudio comparativo

Se ha realizado un estudio comparativo sobre una serie de extracciones respiratorias recibidas por el Centre National de Référence des virus influenzae (Région-Nord), y susceptibles de contener SRAS-CoV. Para ello, se ha extraído el ARN de las extracciones con la ayuda del kit "Qiaamp viral RNA extraction" (Qiagen), y analizado por RT-PCR en tiempo real, por un lado con la ayuda de los pares de cebadores y de sondas de la serie n° 2 en las condiciones descritas anteriormente y, por otro lado, con la ayuda del kit "LightCycler SRAS-CoV quantification kit" comercializado por Roche (referencia 03 604 438). Los resultados se resumen en la tabla VI a continuación. Muestran que 18 de las 26 extracciones son negativas y 5 de las 26 extracciones son positivas para los dos kits, mientras que una extracción es positiva para el único kit Roche y dos para los únicos agentes reactivos N "serie 2". Además, para 3 extracciones (20032701, 20032712, 20032714) las cantidades de ARN detectadas son claramente superiores con los agentes reactivos (sondas y cebadores) de la serie n° 2. Estos resultados indican que los cebadores y las sondas N "serie 2" son más sensibles para la detección del genoma del SRAS-CoV en unas extracciones biológicas que las del kit actualmente disponible.

Tabla VI: Análisis por RT-PCR en tiempo real de los ARN extraídos de una serie de extracciones de 5 pacientes con la ayuda de los pares de cebadores y de sondas de la serie n° 2 (N "serie 2") o del kit "LightCycler SARS-CoV quantification kit" (Roche). El tipo de extracción se indica así como el número de copias del genoma viral medidos en cada uno de los dos ensayos. NEG: RT-PCR negativa.

Extracciones n°	Paciente	Tipo de extracción	KIT ROCHE	N "serie 2"
20033082	K	nasal	NEG	NEG
20033083	K	faríngea	NEG	NEG
20033086	K	nasal	NEG	NEG
20033087	K	faríngea	NEG	NEG
20032802	M	nasal	NEG	NEG
20032803	M	expectoración	NEG	NEG
20032806	M	nasal o faríngea	NEG	NEG
20031746ARN2	C	faríngea	NEG	NEG
20032711	C	nasal o faríngea	39	NEG
20032910	B	nasal	NEG	NEG
20032911	B	faríngea	NEG	NEG
20033356	V	expectoración	NEG	NEG
20033357	V	expectoración	NEG	NEG
20031725	K	asp. endotraqueal	NEG	150
20032657	K	asp. endotraqueal	NEG	NEG
20032698	K	asp. endotraqueal	NEG	NEG
20032720	K	asp. endotraqueal	3	5
20033074	K	heces	115	257
20032701	M	faríngea	443	1676
20032702	M	expectoración	NEG	249
20031747ARN2	C	faríngea	NEG	NEG
20032712	C	desconocida	634	6914
20032714	C	faríngea	17	223
20032800	B	nasal	NEG	NEG
20033353	V	nasal	NEG	NEG
20033384	V	nasal	NEG	NEG

b) Segundo estudio comparativo

Los rendimientos de diferentes métodos de RT-PCR anidada y de RT-PCR en tiempo real se compararon después para 121 extracciones respiratorias de posibles casos de SRAS del hospital francés de Hanoi, Vietnam, realizados entre el 4º y el 17º día después del principio de los síntomas. Entre estas extracciones, 14 se habían encontrado positivas durante un primer ensayo que utiliza el método de RT-PCR anidada que tiene como diana el ORF1b (que codifica la replicasa) tal como se describe inicialmente por el Bernhard Nocht Institute (RT-PCR anidada BNI). Informaciones en relación con este ensayo están disponibles en internet, en la dirección <http://www15.bni-hamburg.de/bni/bni2/neu2/getfile.acgi?area=engl=diagnos-tics&pid=4112>.

Los diferentes ensayos comparados en este estudio son:

- el método de RT-PCR cuantitativa, según la invención, con los cebadores y sonda N "serie 2" descrito anteriormente (columna Light Cycler N),
- el ensayo de RT-PCR anidada que tiene por diana el gen del ARN polimerasa descrito anteriormente, desarrollado por el CDC, el BNI y el Instituto Pasteur (RT-PCR anidada CDC/IP),
- el kit ARTUS de referencia "HPA Corona LC RT-PCR Kit # 5601-02", que es un ensayo de RT-PCR en tiempo real que tiene por diana el gen ORF1b,
- el ensayo de RT-PCR anidada de BNI, que tiene por diana también el gen del ARN polimerasa, mencionado anteriormente.

Los inventores han constatado:

1) una variabilidad inter-ensayo para una misma técnica, relacionada con la degradación de la preparación de ARN durante descongelaciones repetidas, en particular para las muestras que contienen las cantidades de ARN más bajas,

2) una sensibilidad reducida de la RT-PCR anidada CDC/IP con relación a la RT-PCR anidada BNI, y

3) una sensibilidad comparable del ensayo RT-PCR cuantitativa, según la invención, (Light Cycler N) con respecto al ensayo Light Cycler (LC) Artus.

Estos resultados, presentados en la tabla VII a continuación, muestran que el ensayo por RT-PCR cuantitativa, según la invención, constituye un excelente complemento, o una alternativa, a los ensayos actualmente disponibles. En efecto, el coronavirus relacionado al SRAS es un virus emergente, susceptible de evolucionar rápidamente. En particular, el gen del ARN polimerasa del coronavirus relacionado al SRAS, que está en el punto de mira de la mayoría de los ensayos actualmente disponibles, puede recombinar con el de otros coronavirus no relacionados al SRAS. El uso de un ensayo que tiene como diana exclusivamente este gen podría entonces conducir a la obtención de falsos negativos.

El ensayo RT-PCR cuantitativo, según la invención, no tiene por diana la misma región genómica que el kit ARTUS, ya que tiene por diana el gen que codifica la proteína N. Realizando un ensayo de diagnóstico que tiene por diana dos genes diferentes del coronavirus relacionado con el SRAS, se puede por lo tanto esperar librarse de resultados de tipo falsos positivos que podrían deberse a la evolución genética del virus.

Además, parece particularmente ventajoso tener por diana el gen de la proteína de nucleocápside, ya que es muy estable, debido a la fuerte presión de selección relacionada con las tensiones estructurales elevadas que conciernen a esta proteína.

Tabla VII: Comparación de diferentes métodos de análisis por amplificación génica, a partir de 121 extracciones de posibles casos de SRAS del hospital francés de Hanoi, Vietnam (epidemia del 2003)

N° CNR	Tipo de extracción (1)	Día de extracción	Paciente	RT-PCR anidada CDC/IP	RT-PCR anidada BNI	kit Light Cycler Artus	Light Cycler N (IP)
107 extracciones	N y P			Negativa	Negativa	Negativa	Negativa
032529	P	10	NHB	Negativa	Positiva	Negativa	Negativa
032530	N	10	NHB	Positiva	Positiva	3,10E+01	4,20E+01
032531	P	7	LP	Positiva	Positiva	7,70E+00	3,10E+00
032534	N	15	BND	Positiva	Positiva	1,60E+00	Negativa
032600	P	4	NHH	Negativa	Positiva	Negativa	1,30E+02
032612	P	17	NTS	Negativa	Positiva	Negativa	Negativa
032688	P	9	BTX	Positiva	Positiva	Negativa	Negativa
032689	N	4	NVH	Positiva	Positiva	1,20E+01	2,30E+02
032690	P	4	NVH	Negativa	Positiva	1,60E+00	Negativa
032727	P	8	NVH	Positiva	Positiva	2,30E+02	4,00E+02
032728	N	8	NVH	Positiva	Positiva	1,10E+03	1,60E+04
032729	P	14	NHB	Positiva	Positiva	5,90E+00	3,40E+01
032730	N	14	NHB	Positiva	Positiva	1,30E+02	4,80E+02
032741	P	8	NHH	Positiva	Positiva	2,10E+02	1,30E+02
Positivas				10	14	10	9
Fracción detectada de las 14 Positivas				71,4%	100,0%	71,4%	64,3%

(1) P= hisopo faríngeo
N= hisopo nasal

Ejemplo 9: Obtención y caracterización de anticuerpos monoclonales dirigidos contra la proteína N

Se inmunizaron ratones Balb C con la ayuda de la proteína N recombinante purificada y se fusionaron sus células esplénicas con un mieloma murino apropiado, según las técnicas de Kohler y Milstein.

Se preseleccionaron diez y nueve hibridomas secretores de anticuerpos anti N y se han preseleccionado sus inmunorreactividades precisadas. Estos anticuerpos reconocen bien la proteína N recombinante (en ELISA) con unas intensidades variables, así como la proteína viral, natural N en ELISA y/o transferencia western. Las figuras 18 a 20 muestran los resultados de estos ensayos para 15 de estos 19 anticuerpos monoclonales.

Se han sub-clonado los clones 12, 17, 28, 57, 72, 76, 86, 87, 98, 103, 146, 156, 166, 170, 199, 212, 218, 219 y 222, fuertemente reactivos. Los estudios de especificidad se prosiguieron con las herramientas apropiadas a fin de precisar los epítomos reconocidos, y verificar la ausencia de reactividad frente a otros coronavirus humanos y de algunos virus respiratorios.

Los estudios de cartografía (*mapping*) epitópica (realizados sobre membrana "spot", con la ayuda de péptidos superpuestos de 15 aa) y los estudios suplementarios realizados sobre la proteína N natural en transferencia western revelaron la existencia de 4 grupos de anticuerpos monoclonales:

1° Anticuerpos monoclonales específicos de un epítipo lineal mayor en posición N-ter (75-81, secuencia: INTNSVP). El representante de este grupo es el anticuerpo 156. El hibridoma que produce este anticuerpo se depositó en la Collection Nationale de Cultures de Microorganismes (CNCM) del Instituto Pasteur (Paris, Francia) el 1 de diciembre de 2004, bajo el número I-3331. Este mismo epítipo es también reconocido por un suero de conejo (policlonal anti N) obtenido mediante inmunización clásica con la ayuda de esta misma proteína N.

2° Anticuerpos monoclonales específicos de un epítipo lineal mayor situado en posición central (posición 217-224, secuencia: ETALALL); los representantes de este grupo son los anticuerpos monoclonales 87 y 166. El hibridoma que produce el anticuerpo 87 se depositó en la CNCM el 1 de diciembre de 2004, bajo el número I-3328.

3° Anticuerpos monoclonales específicos de un epítipo lineal mayor situado en posición C-terminal (posición 403-408, secuencia: DFFRQL), los representantes de este grupo son los anticuerpos 28, 57 y 143. El hibridoma que produce el anticuerpo 57 se depositó en la CNCM el 1 de diciembre de 2004, bajo el número I-3330.

4° Anticuerpos monoclonales específicos de un epítipo discontinuo, conformacional. Este grupo de anticuerpos no reconoce ninguno de los péptidos que recubren la secuencia de la proteína N, pero reaccionan fuertemente sobre la proteína natural no desnaturalizada. El representante de este último grupo es el anticuerpo 86. El hibridoma que produce este anticuerpo se depositó en la CNCM el 1 de diciembre de 2004, bajo el número I-3329.

La tabla VIII siguiente resume los resultados de cartografía epitópica obtenidos:

Tabla VIII: Cartografía epitópica de los anticuerpos monoclonales

Anticuerpos	Epítipo	Posición	Región
28	DFSRQL Q	403 ... 408	C - Ter.
143	DFSRQL Q		
76	DFSRQL Q		
57	DFSRQL Q		
	FFGMS RI	315 ... 319	
146	LPQRQ	383 ... 387	
166	ETALALLL	217 ... 224	central
87	ETALALL	217 ... 224	
156	INTNSGP	75 ... 81	N - Ter.
86	Conformacional		
212	Conformacional		
170	Conformacional		

Además, como se ilustra en particular en las figuras 18 a 19, estos anticuerpos no presentan reactividad en ELISA y/o en WB, frente a la proteína N del coronavirus humano 229 E.

Ejemplo 10: Combinaciones de anticuerpos monoclonales para el desarrollo de un ensayo de inmunocaptura sensible y específica del antígeno viral N en el suero o los fluidos biológicos de los pacientes contaminados por el virus SRAS CoV

Los anticuerpos listados a continuación se seleccionaron debido a sus propiedades bien particulares para un estudio suplementario de captura y detección de la proteína viral N, en el suero de los sujetos o pacientes.

Estos anticuerpos se produjeron en ascitis en ratones, purificados mediante cromatografía de afinidad y utilizados solos o en combinación como anticuerpos de captura, y como anticuerpos señal.

Listado de los anticuerpos seleccionados para este propósito:

- Acm anti región C-ter (n° 28, 57, 143)

- Acm anti región central (n° 87, 166)
- Acm anti región N-ter (n° 156)
- Acm anti epítipo discontinuo, conformacional (86)

1) Preparación de los agentes reactivos:

a) Placas ELISA de inmunocaptura

Las placas son sensibilizadas con las soluciones de anticuerpos a 5 µg/ml de tampón carbonato 0,1 M, pH 9,6. Las soluciones (monovalentes o plurivalentes) son depositadas en un volumen de 100 µl en los pocillos, y se incuban durante una noche a temperatura ambiente. Estas placas se lavan después en tampón PBS (10 mM pH 7,4 adicionado del 0,1% de Tween 20) y después se saturan con una solución de PBS adicionada del 0,3% de BSA y del 5% de sacarosa). Las placas se secan después y se envasan en una bolsa en presencia de un desecador. Están listas para su empleo.

b) Conjugados

Los anticuerpos purificados se acoplaron con la peroxidasa según el protocolo de Nakane (Nakane y otros - 1974, J. of Histo and cytochemistry, vol. 22, p 1084-1091) en una relación de una molécula de IgG para 3 moléculas de peroxidasa. Estos conjugados se purificaron mediante cromatografía de exclusión y se conservaron concentrados (concentración comprendida entre 1 a 2 mg/ml) en presencia del 50% de glicerol y a -20°C. Están diluidos para su utilización en los ensayos para la concentración final de 1 ó 2 µg/ml en tampón PBS (pH 7,4) adicionado del 1% de BSA.

c) Otros agentes reactivos

- Sueros humanos negativos para todos los marcadores séricos de los virus HIV, HBV, HCV y THLV
- Conjunto (pool) de sueros humanos negativos con adición de 0,5% de Tritón X 100
- Ag viral inactivo: sobrenadante de cultivo viral inactivo mediante irradiación e inactivación verificada después del cultivo sobre células sensibles - título de la suspensión antes de la inactivación de aproximadamente 10^7 partículas infecciosas por ml o también aproximadamente 5×10^9 partículas virales físicas por ml de antígeno
- Las muestras de Ag diluidas en suero humano negativo: estas muestras se prepararon mediante dilución al 1:100 y después por dilución en serie de razón 5.

Estas muestras no infecciosas mimetizan unas muestras humanas que esperan contener concentraciones bajas a muy bajas de nucleoproteína viral N. Tales muestras no son accesibles para los trabajos rutinarios

- Solución de lavado R2, solución de revelado TMB R8, cromógeno R9 y solución de detención R10, son los agentes reactivos genéricos comercializados por Bio-Rad en sus estuches ELISA (ej.: estuche Platelia Pylori ref. 72778).

2) Modo de realización

Las muestras de sueros humanos sobrecargados en Ag viral desactivado son distribuidas a razón de 100 µl por pozo, directamente en las placas sensibilizadas, listas para su empleo y después incubados durante 1 hora a 37°C (incubación Bio-Rad IPS).

El material no retenido por la fase sólida se elimina mediante 3 lavados (lavado con solución R2 diluida, lavador automático LP 35).

Los conjugados apropiados, diluidos a la concentración final de 1 ó 2 µg/ml son distribuidos a razón de 100 µl por pozo y las placas se incuban de nuevo durante una hora a 37°C (incubación IPS).

El exceso de conjugado se elimina mediante 4 lavados sucesivos (solución R2 diluida - lavador LP 35).

La presencia de conjugado fijado sobre las placas se revela después de la adición de 100 µl de solución de revelado preparada antes del uso (1 ml de R9 y 10 ml de R8) y después de la incubación durante 30 minutos, a temperatura ambiente y protegido de la luz.

La reacción enzimática se bloquea finalmente mediante adición de 100 µl de agente reactivo R10 (H₂SO₄ 1N) en todas los pozos.

La lectura se efectúa con la ayuda de un lector de microplacas apropiado de doble longitud de onda (450/620 nm).

Los resultados pueden ser interpretados utilizando como valor límite provisional la media de al menos dos controles negativos multiplicado por un factor 2 o también la media de 100 sueros negativos adicionada de un incremento que corresponde a 6 SD (Desviación estándar calculada sobre las 100 mediciones individuales).

3) Resultados

Diferentes combinaciones anticuerpos de captura y anticuerpos señal se ensayaron basándose en las propiedades de los anticuerpos seleccionados, y evitando las combinaciones de anticuerpos específico de los mismos epítomos en fase sólida y conjugados

Los mejores resultados se obtuvieron con las 4 combinaciones listadas a continuación. Estos resultados son reproducidos en la tabla IX a continuación.

1. Combinación F/28

Fase sólida (Acm 166 + 87 región central): conjugado anticuerpo 28 (C-ter)

2. Combinación G/28

Fase sólida (Acm 86 - epítipo conformacional): conjugado anticuerpo 28 (C-ter)

3. Combinación H/28

Fase sólida (Acms 86, 166 y 87 región central y epítipo conformacional): conjugado anticuerpo 28 (C-ter)

4. Combinación H/28 + 87

Fase sólida (Acms 86, 166 y 87 región central y epítipo conformacional): conjugado mixto anticuerpo 28 (C-ter) y 87 (central)

5. Combinación G/87

Fase sólida (Acm 86 - epítipo conformacional): conjugado anticuerpo 87 (región central)

Las 4 primeras combinaciones presentan unos rendimientos equivalentes y reproducibles, superiores a las otras combinaciones utilizadas (por ejemplo la combinación G/87). Por supuesto, en estas combinaciones, un anticuerpo monoclonal puede ser sustituido con otro anticuerpo que reconoce el mismo epítipo. Así, se pueden citar las variantes siguientes:

6. Variante de la combinación F/28

Fase sólida (Acm 87 únicamente): conjugado anticuerpo 57 (C-ter)

7. Variante de la combinación G/28

Fase sólida (Acm 86 - epítipo conformacional): conjugado anticuerpo 57 (C-ter)

8. Variante de la combinación H/28

Fase sólida (Acms 86 y 87 región central y epítipo conformacional): conjugado anticuerpo 57 (C-ter)

9. Variante de la combinación H/28 + 87

Fase sólida (Acms 86 y 87 región central y epítipo conformacional): conjugado mixto anticuerpo 57 (C-ter) y 87 (central)

Tabla IX: Control de la inmunorreactividad de los Acm anti-nucleoproteínas SRAS CoV: densidades ópticas medidas con cada combinación de anticuerpos, en función de las diluciones del antígeno viral inactivo.

Nº	Dilución	F/28	G/28	G/87	H/28	H/28+87
0	1/100	5	5	3,495	3,900	5
1	1/500	3,795	3,814	1,379	3,702	3,804
2	½ 500	2,815	2,950	0,275	3,268	2,680
3	1/12500	0,987	1,038	0,135	1,374	0,865
4	1/62500	0,404	0,348	0,125	0,480	0,328
5	1/312500	0,285	0,211	0,123	0,240	0,215
6	Control	0,210	0,200	0,098	0,186	0,156
7	Control	0,269	0,153	0,104	0,193	0,202

El límite de detección de estos 4 ensayos experimentales corresponde a la dilución de antígeno en suero negativo 1: 62500. Una extrapolación rápida deja suponer la detección de menos de 10^3 partículas por ml de sueros.

De este estudio, destaca que los anticuerpos más apropiados para la captura de la nucleoproteína viral nativa son los anticuerpos específicos de la región central y/o de un epítipo conformacional, siendo uno y otro anticuerpos seleccionados también por su fuerte afinidad para el antígeno nativo.

Habiendo determinado los mejores anticuerpos para la composición de la fase sólida, los anticuerpos para conservar la prioridad para la detección de antígenos fijados sobre la fase sólida son los anticuerpos complementarios específicos de un epítipo dominante en la región C-ter. El empleo de cualquier otro anticuerpo complementario, pero específico de los epítopos localizados en la región N-ter de la proteína, conduce a resultados medios o mediocres.

Ejemplo 1.1: Sistemas de expresión eucariotas de la proteína de espícula (S) del coronavirus asociado al SRAS (SRAS-CoV)

1) Optimización de las condiciones de expresión de la S del SRAS-CoV en células de mamíferos.

Las condiciones de expresión transitoria de la proteína de espícula (S) del SRAS-CoV se optimizaron en células de mamíferos (293T, VeroE6).

Para ello, se amplificó un fragmento de ADN que contiene el ADNc de la S del SRAS-CoV mediante PCR con la ayuda de los oligonucleótidos 5'-ATAGGATCCA CCATGTTTAT TTTCTTATTA TTTCTTACTC TCACT-3' y 5'-ATACTCGAGTT ATGTGTAATG TAATTTGACA CCCTTG-3' a partir del plásmido pSRAS-S (C.N.C.M. n° 1-3059) y después se insertó en los sitios BamH1 y Xho1 del plásmido pTRIPΔU3-CMV que contiene un vector lentiviral TRIP (Sirven, 2001, Mol. Ther., 3, 438-448) para obtener el plásmido pTRIP-S. El fragmento BamH1 y Xho1 que contiene el ADNc de la S se sub-clonó después entre los BamH1 y Xho1 del plásmido de expresión eucariota pcDNA3.1(+) (Clontech) para obtener el plásmido pcDNA-S. El fragmento Nhe1 y Xho1 que contiene el ADNc de la S se sub-clonó después entre los sitios correspondientes del plásmido de expresión pCI (Promega) para obtener el plásmido pCI-S. Las secuencias WPRE del virus de la hepatitis de la marmota ("Woodchuck Hepatitis Virus posttranscriptional regulatory element") y las secuencias CTE ("constitutive transport element") del retrovirus simio de Mason-Pfizer se insertaron en cada uno de los dos plásmidos pcDNA-S y pCI-S entre los sitios Xho1 y Xba1 para obtener respectivamente los plásmidos ADNpc-S-CTE, ADNpc-S-WPRE, pCI-S-CTE y pCI-S-WPRE (figura 21). El plásmido pCI-S-WPRE se depositó en la CNCM, el 22 de noviembre de 2004, bajo el número 1-3323. Todos los insertos se secuenciaron con la ayuda de un kit BigDye Terminator v1.1 (Applied Biosystems) y de un secuenciador automático ABI377.

Se buscó la capacidad de los plásmidos construidos para dirigir la expresión de la S del SRAS-CoV en células de mamíferos después de la transfección de células VeroE6 (figura 22). En este experimento, se han transfectado monocapas de 5×10^5 células VeroE6 en cajas de Petri de 35 mm con 2 μ g de los plásmidos ADNpc (a título de control), ADNpc-S, pCI y pCI-S y 6 μ l de agente reactivo Eugene6 según las indicaciones del fabricante (Roche). Después de 48 horas de incubación a 37°C y bajo 5% de CO₂, se prepararon unos extractos celulares en tampón de depósito según Laemmli, separados sobre un gel SDS al 8% de poliacrilamida, y después transferidos sobre una membrana de PVDF (BioRad). La detección de esta inmunohuella («western blot») se realizó con la ayuda de un suero policlonal de conejo anti-S (suero inmune del conejo P11135: véase el ejemplo 4 anterior) y de anticuerpos policlonales de asno dirigidos contra los IgG de conejo y acoplados con la peroxidasa (NA934V, Amersham). Los anticuerpos fijados se revelaron mediante luminescencia con la ayuda del kit ECL+ (Amersham) y de películas de autorradiografía Hyperfilm MP (Amersham).

Este experimento (figura 22) muestra que el plásmido pcDNA-S no permite dirigir la expresión de la S del SRAS-CoV a niveles detectables mientras que el plásmido pCI-S permite una expresión baja, próxima del límite de detección, que puede ser demostrado cuando la película está sobre-expuesta. Unos resultados similares se obtuvieron cuando la expresión de la S se buscó mediante inmunofluorescencia (datos no mostrados). Esta imposibilidad de detectar una expresión eficaz de la S no se puede imputar a las técnicas de detección utilizadas ya que la proteína S puede ser puesta en evidencia al tamaño esperado (180 kDa) en un extracto de células infectadas por el SRAS-CoV o en un extracto de células VeroE6 infectadas por el virus recombinante de la vacuna antivariólica VV-TF7.3 y transfectadas por el plásmido pcDNA-S. En este último experimento, el virus VV-TF7.3 expresa el ARN polimerasa del fago T7, y permite la transcripción citoplásmica de un ARN sin capuchón susceptible de ser traducido eficazmente. Este experimento sugiere que los defectos de expresión descritos anteriormente se deben a una incapacidad intrínseca del ADNc de la S para ser expresado eficazmente cuando la etapa de transcripción en ARN mensajero se realiza a nivel nuclear.

En un segundo experimento, el efecto de las señales CTE y WPRE sobre la expresión de la S se ha buscado después de la transfección de células VeroE6 (figura 23A) y 293T (figura 23B) y según un protocolo similar al descrito anteriormente. Mientras que la expresión de la S no se puede demostrar después de la transfección de los

plásmidos ADNpc-S-CTE y ADNpc-S-WPRE derivados de ADNpc-S, la inserción de las señales WPRE y CTE mejora fuertemente la expresión de la S en el contexto del plásmido de expresión pCI-S.

Para precisar este resultado, se realizó una segunda serie de experimentos, en la que la inmunohuella se revela de manera cuantitativa por luminescencia y adquisición sobre un dispositivo de imágenes digitales (FluorS, BioRad). El análisis de los resultados obtenidos con el programa QuantityOne v4.2.3 (BioRad) muestra que las secuencias WPRE y CTE aumentan respectivamente la expresión de la S de un factor 20 a 42 y 10 a 26 en células Vero E6 (tabla X). En células 293T (tabla X), el efecto de la secuencia CTE es más moderado (4 a 5 veces), mientras que el de la secuencia WPRE sigue importante (13 a 28 veces).

Tabla X: Análisis cuantitativo del efecto de las señales CTE y WPRE sobre la expresión de la S del SRAS-CoV:

Se prepararon extractos celulares 48 horas después de la transfección de células VeroE6 o 293T por los plásmidos pCI, pCI-S, pCI-S-CTE y pCI-S-WPRE, y se analizaron mediante transferencia western como se describe en la leyenda de la figura 22. La transferencia western se revela mediante luminescencia (ECL+, Amersham) y adquisición sobre un dispositivo de imágenes digitales (FluorS, BioRad). Los niveles de expresión son indicados en función de una escala arbitraria en la que el valor de 1 representa el nivel medido después de la transfección del plásmido pCI-S. Se realizaron dos experimentos independientes para cada uno de los dos tipos celulares. En el experimento 1, sobre células VeroE6, las transfecciones se realizaron en duplicado y los resultados son indicados en forma de media y desviaciones estándares de los niveles de expresión medidos.

Plásmido	Célula	exp. 1	exp. 2
PCI	VeroE6	0,0	0,0
pCI-S	VeroE6	1,0 ± 0,1	1,0
pCI-S-CTE	VeroE6	9,8 ± 0,9	26,4
pCI-S-WPRE	VeroE6	20,1 ± 2,0	42,3
PCI	293T	0,0	0,0
PCI-S	293T	1,0	1,0
PCI-S-CTE	293T	4,6	4,0
PCI-S-WPRE	293T	27,6	12,8

En resumen, el conjunto de estos resultados muestra que la expresión en unas células de mamíferos del ADNc de la S del SRAS-CoV, bajo la dependencia de secuencias promotoras del ARN polimerasa II, requiere, para ser eficaz, la presencia de una señal de empalme así como una u otra de las secuencias WPRE y CTE.

2) Obtención de líneas estables que permiten la expresión de la S del SRAS-CoV

El ADNc de la proteína S del SRAS-CoV se clonó en forma de un fragmento BamH1-Xho1 en el plásmido pTRIP)AU3-CMV que contiene un vector lentiviral defectivo TRIP a ADN flap central (Sirven y otros, 2001, Mol. Ther., 3: 438-448), para obtener el plásmido pTRIP-S (figura 24). La co-transfección transitoria según Zennou y otros (2000, Cell, 101: 173-185) de este plásmido, de un plásmido de encapsidación (p8.2), y de un plásmido de expresión de la glicoproteína de envoltura G del VSV (pHCMV-G) en unas células 293T, permitió la preparación de pseudopartículas retrovíricas que contienen el vector TRIP-S, y pseudotipadas por la proteína de envoltura G. Estos vectores TRIP-S pseudotipados se utilizaron para transducir unas células 293T y FRhK-4: no se ha podido demostrar ninguna expresión de la proteína S, ni ha podido ser puesto en evidencia mediante transferencia western e inmunofluorescencia en las células transducidas (datos no presentados).

Los casetes de expresión óptimos constituidos del promotor inmediato/precoz del virus CMV, de una señal de corte y empalme, del ADNc de la S y de una u otra de las señales post-transcripcionales WPRE o CTE, descritas anteriormente, se sustituyeron entonces por el casete EF1 α -EGFP del vector de expresión lentiviral defectivo a ADN FLAP central TRIP)AU3-EF1 α (Sirven y otros, 2001, Mol. Ther., 3:438-448) (figura 25). Estas sustituciones se realizaron mediante una serie de sub-clonaciones sucesivas de los casetes de expresión de S escindidos de los plásmidos pCT-S-CTE (BglII-Ap1) o respectivamente pCI-S-WPRE (BglII-Sal1), y después se insertan entre los sitios Mlu1 y Kpn1, o respectivamente Mlu1 y Xho1, del plásmido TRIP)AU3-EF1 α para obtener los plásmidos pTRIP-SD/SA-S-CTE y pTRIP-SD/SA-S-WPRE, depositados en la CNCM, el 1 de diciembre de 2004, bajo los números I-3336 y I-3334, respectivamente. Unos vectores pseudotipados se produjeron según Zennou y otros (2000, Cell, 101: 173-185) y utilizados para transducir unas células 293T (10000 células) y FRhK-4 (15000 células) según una

serie de 5 ciclos sucesivos de transducción con una cantidad de vector que corresponde a 25 ng (TRIP-SD/SA-S-CTE) o 22 ng TRIP-SD/SA-S-WPRE) de p24 por ciclo.

Las células transducidas se clonaron por dilución límite y se analizó una serie de clones para la expresión de la S del SRAS-CoV cualitativamente por inmunofluorescencia (datos no mostrados), y después cuantitativamente por transferencia western (figura 25) con la ayuda de un suero policlonal de conejo anti-S. Los resultados presentados en la figura 25 muestran que los clones 2 y 15 de células FRhK4-s-CTE transducidas por TRIP-SD/SA-S-CTE y los clones 4, 9 y 12 de células FRhK4-S-WPRE transducidas por TRIP-SD/SA-S-WPRE, permiten la expresión de la S del SRAS-CoV a niveles respectivamente bajos y moderados si se les comparan con los que se pueden observar durante la infección por el SRAS-CoV.

En resumen, los vectores TRIP-SD/SA-S-CTE y TRIP-SD/SA-S-WPRE permiten la obtención de clones estables de células FRhK-4 y, de manera similar, 293T que expresa la S del SRAS-CoV, mientras que los ensayos realizados con el vector "de base" TRIP-S siguen infructuosos, lo que demuestra la necesidad de una señal de empalme, así como una u otra de las secuencias CTE y WPRE para la obtención de clones celulares estables que expresan la proteína S.

Además, estas modificaciones del vector TRIP (inserción de una señal de corte y empalme empalme post-transcripcional como CTE y WPRE), podrían revelarse interesantes para mejorar la expresión de otros ADNc diferente de aquel de la S.

3) Obtención de líneas estables que permiten la expresión de una forma soluble de la S del SRAS-CoV. Purificación de este antígeno recombinante

Un ADNc que codifica una forma soluble de la proteína S (Ssol) se obtuvo fusionando las secuencias que codifican para el ectodominio de la proteína (aminoácidos 1 a 1193) a las de una etiqueta (FLAG: DYKDDDDK) a través de un enlazador BspE1, que codifica el dipéptido SG. Prácticamente, para obtener el plásmido ADNpc-Ssol, se amplificó un fragmento de ADN que codifica el ectodominio de la S del SRAS-CoV por PCR con la ayuda de los oligonucleótidos 5'-ATAGGATCCA CCATGTTTAT TTTCTTATTA TTTCTTACTC TCACT-3' y 5'-ACCTCCGGAT TTAATATATT GCTCATATTT TCCCAA- 3' a partir del plásmido pcDNA-S, y después se insertó entre los sitios únicos BamH1 y BspE1 únicos de un plásmido de expresión eucariota ADNpc3.1 (+) (Clontech) modificado, que contiene entre sus sitios BamH1 y Xho1 la secuencia de la etiqueta FLAG:

```
// GGATCC ...nnn... TCC GGA GAT TAT AAA GAT GAC GAC GAT AAA TAA
      BamH1           S   G   D   Y   K   D   D   D   D   K   ter
CTCGAG //
Xho1
```

Los fragmentos Nhe1-Xho1 y BamH1-Xho1, que contiene el ADNc de la forma soluble de la S, se extirparon después del plásmido pcDNA-Ssol, y se sub-clonaron entre los sitios correspondientes del plásmido pTRIP-SD/SA-S-CTE y del plásmido pTRIP-SD/SA-S-WPRE, respectivamente, para obtener los plásmidos pTRIP-SD/SA-Ssol-CTE y pTRIP-SD/SA-Ssol-WPRE, depositados en la CNCM, el 1 de diciembre de 2004, bajo los números 1-3337 e 1-3335, respectivamente.

Se produjeron unos vectores pseudotipos según Zennou y otros (2000, Cell, 101: 173-185) y se utilizaron para transducir unas células FRhK-4 (15000 células) según una serie de 5 ciclos sucesivos de transducción (15000 células) con una cantidad de vector que corresponde a 24 ng (TRIP-SD/SA-Ssol-CTE) o 40 ng TRIP-SD/SA-Ssol-WPRE) de p24 por ciclo. Las células transducidas se clonaron por dilución límite y una serie de 16 clones transducidos mediante TRIP-SD/SA-Ssol-CTE y de 15 clones mediante TRIP-SD/SA-Ssol-WPRE se analizaron para la expresión del polipéptido Ssol mediante transferencia western, revelado por un anticuerpo monoclonal anti-FLAG (figura 26 y datos no mostrados), así como mediante un ELISA-captura específico del polipéptido Ssol que se realizó con este objetivo (tabla XI y datos no mostrados). Una parte del proceso de selección de los mejores clones secretores se muestra en la figura 26. El ELISA-captura se basa en el uso de fases sólidas recubiertas de anticuerpos policlonales de conejos inmunizados por SRAS-CoV purificado e inactivado. Estas fases sólidas permiten la captura del polipéptido Ssol secretado en los sobrenadantes celulares, cuya presencia se revela después mediante una serie de etapas que implican sucesivamente la fijación de un anticuerpo monoclonal anti-FLAG (M2, SIGMA), de anticuerpos biotinilados policlonales de conejo anti-IgG(H+L), de ratón (Jackson), y de un conjugado streptavidina-peroxidasa (Amersham), y después la adición de cromógeno y de sustrato (TMB + H₂O₂, KPL).

Tabla XI: análisis de la expresión de polipéptido Ssol por unas líneas celulares transducidas por los vectores lentivíricos TRIP-SD/SA-Ssol-WPRE y TRIP-SD/SA-Ssol-CTE. La secreción del polipéptido Ssol se buscó en el sobrenadante de una serie de clones celulares aislados después de la transducción de células FRhK-4 por los vectores lentivíricos TRIP-SD/SA-Ssol-WPRE y TRIP-SD/SA-Ssol-CTE. Los sobrenadantes diluidos al 1/50 se analizaron mediante un ensayo ELISA-captura específico de la S del SRAS-CoV.

Vector	Clon	DO (450 nm)
Control	-	0,031
TRIP-SD/SA-Ssol-CTE	CTE2	0,547
	CTE3	0,668
	CTE9	0,171
	CTE12	0,208
	CTE13	0,133
TRIP-SD/SA-Ssol-WPRE	WPRE1	0,061
	WPRE10	0,134

La línea celular que segrega las cantidades más elevadas de polipéptido Ssol en el sobrenadante de cultivo es la línea FRhK4-Ssol-CTE3. Se ha sometido a una segunda serie de 5 ciclos de transducción por el vector TRIP-SD/SA-Ssol-CTE, en condiciones similares a las descritas anteriormente, y después se ha clonado. El sub-clon que segrega las cantidades más elevadas de Ssol se ha seleccionado mediante una combinación de análisis por transferencia western, y ELISA-captura: se trata del sub-clon FRhK4-Ssol-30, que se depositó en la CNCM, el 22 de noviembre de 2004, bajo el número I-3325.

La línea FRhK4-Ssol-30 permite la producción y la purificación en cantidad del polipéptido recombinante Ssol. En un experimento del tipo en el que las condiciones experimentales de crecimiento, de producción y de purificación se optimizaron, las células de la línea FRhK4-Ssol-30 son inoculadas en un medio de cultivo estándar (DMEM sin piruvato que contiene 4,5g/l de glucosa y suplementado con 5% de SVF, 100 U/ml de penicilina y 100 µg/ml de estreptomicina) en forma de una monocapa sub-confluente (1 millón de células para cada 100 cm² en 20 ml de medio). A confluencia, el medio estándar se sustituye con el medio de secreción en el que la cantidad de SVF se reduce al 0,5% y la cantidad de medio se reduce a 16 ml para cada 100 cm². El sobrenadante de cultivo se extrae después de 4 a 5 días de incubación a 35°C y bajo el 5% de CO₂. El polipéptido recombinante Ssol se purifica a partir del sobrenadante mediante la cadena de etapas de filtración sobre membrana de poliétersulfona (PES) de 0,1 µm, de concentración por ultra-filtración sobre una membrana de PES de punto de escisión 50 kD, de cromatografía de afinidad sobre plantilla anti-FLAG con elución mediante una solución de péptido FLAG (DYKDDDDK) a 100 µg/ml en TBS (Tris 50 mM pH 7,4, 150 mMNaCl), y después de cromatografía de gel de filtración en TBS sobre bolas de sephadex G-75 (Pharmacia). La concentración del polipéptido recombinante Ssol purificado se determinó mediante el ensayo micro-BCA (Pierce), y después se analizaron sus características bioquímicas.

El análisis por gel SDS al 8% de acrilamida teñida con nitrato de plata pone de manifiesto un polipéptido mayoritario cuya masa molecular es de aproximadamente 180 kD y cuyo grado de pureza puede ser evaluado al 98% (figura 27A). Mediante espectrometría de masas SELDI-TOF (Cyphergeren), se ponen en evidencia dos picos principales: corresponden a formas simple y doblemente cargadas de un polipéptido mayoritario, cuya masa molecular está así determinada a 182,6 ± 3,7 kD (figuras 27B y C). Después de la transferencia sobre membrana Prosorb y del aclarado en TFA 0,1%, el extremo N-terminal del polipéptido Ssol se secuenció en fase líquida mediante degradación de Edman sobre 5 residuos (ABI494, Applied Biosystems), y determinada como siendo SDLDR (figura 27D). Esto demuestra que el péptido señal localizado en el extremo N-terminal de la proteína S del SRAS-CoV, compuesto de las aa 1 a 13 (MFIFLLFLTLTSG) según un análisis realizado con el programa signalP v2.0 (Nielsen y otros, 1997, Protein Engineering, 10: 1-6), está escindido del polipéptido Ssol maduro. El polipéptido recombinante Ssol está por lo tanto constituido de los aminoácidos 14 a 1193 de la proteína S del SRAS-CoV fusionados en C-terminal con una secuencia SGDYKDDDDK que contiene la secuencia de la etiqueta FLAG (subrayada). La diferencia entre la masa molar teórica del polipéptido Ssol desnudo (132,0 kD) y la masa molar real del polipéptido maduro (182,6 kD) sugiere que el polipéptido Ssol está glicosilado.

Una preparación de polipéptido Ssol purificado, y cuya concentración proteica se determinó mediante un ensayo micro-BCA, permite realizar una escala estándar para medir, con la ayuda del ensayo ELISA-captura descrito anteriormente, las concentraciones de Ssol presente en los sobrenadantes de cultivo y volver a considerar las características de las líneas secretoras. Según este ensayo, la línea FRhK4-Ssol-CT3 segrega 4 a 6 µg/ml de polipéptido Ssol mientras que la línea FRhK4-Ssol-30 segrega 9 a 13 µg/ml de Ssol después de 4 a 5 días de cultivo a confluencia. Además, el esquema de purificación presentado anteriormente permite, rutinariamente, purificar de 1 a 2 mg de polipéptido Ssol por litro de sobrenadante de cultivo.

Ejemplo 12: Inmunización génica que implica a la proteína de espícula (S) del coronavirus asociado al SRAS (SRAS-CoV)

El efecto de una señal de empalme y de las señales post-transcripcionales WPRE y CTE, se analizó después de la inmunización génica de ratones BALB/c (figura 28).

Para ello, se inmunizaron unos BALB/c a intervalos de 4 semanas por inyección en el tibialis anterior de una solución salina de 50 µg de ADN plasmídico de pcDNA-S y pCI-S así como, a título de control, mediante 50 µg de ADN plasmídico de ADNpc-N (que dirige la expresión de la N del SRAS-CoV), o de pCI-HA (que dirige la expresión de la HA del virus gripal A/PR/8/34), y se recogen los sueros inmunes 3 semanas después de la 2ª inyección. La

presencia de anticuerpos dirigidos contra la S del SRAS-CoV se buscó mediante ELISA indirecto utilizando como antígeno un lisado de células VeroE6 infectadas por el SRAS-CoV y a título de control, un lisado de VeroE6 no infectadas. Los títulos (TI) en anticuerpos anti-SRAS-CoV se calculan como la inversa de la dilución que produce una DO específica de 0,5 (diferencia entre DO medida sobre lisado de células infectadas y DO medida sobre lisado de células no infectadas) después del revelado mediante un anticuerpo policlonal anti-IgG de ratón acoplado a la peroxidasa (NA931V, Amersham) y de TMB adicionado de H₂O₂ (KPL) (figura 28A).

En estas condiciones, el plásmido de expresión pcDNA-S permite sólo la inducción de bajos títulos de anticuerpos dirigidos contra la S del SRAS-CoV en 3 ratones de 6 ($\text{LOG}_{10}(\text{TI})=1,9\pm0,6$), mientras que el plásmido ADNpc-N permite la inducción de anticuerpos anti-N con títulos elevados ($\text{LOG}_{10}(\text{TI})=3,9\pm0,3$) en todos los animales, y los plásmidos controles (pCI, pCI-HA) no conducen a ningún anticuerpo detectable ($\text{LOG}_{10}(\text{TI})<1,7$). El plásmido pCI-S provisto de una señal de corte y empalme permite la inducción de anticuerpos con títulos elevados ($\text{LOG}_{10}(\text{TI})=3,7\pm0,2$), que son aproximadamente 60 veces superiores a los observados después de la inyección del plásmido pcDNA-S ($p<10^{-5}$).

La eficacia de las señales post-transcripcionales se estudió realizando un estudio dosis-respuesta de los títulos en anticuerpos anti-S inducidos en ratón BALB/c en función de la cantidad de ADN plasmídico utilizado como inmunógeno (2 µg, 10 µg y 50 µg). Este estudio (figura 28B) demuestra que la señal post-transcripcional WPRE mejora fuertemente la eficacia de la inmunización génica cuando se utilizan bajas dosis de ADN ($p<10^{-5}$ para una dosis de 2 µg de ADN y $p<10^{-2}$ para una dosis de 10 µg), mientras que el efecto de la señal CTE sigue siendo marginal ($p=0,34$ para una dosis de 2 µg de ADN).

Finalmente, los anticuerpos inducidos en el ratón después de la inmunización génica neutralizan la infectividad del SRAS-CoV *in vitro* (figuras 29A y 29B) a títulos que están en relación con los títulos medidos por ELISA.

En resumen, el uso de una señal de empalme y de la señal post-transcripcional WPRE del virus de la hepatitis de la marmota mejora de manera considerable la inducción de anticuerpos neutralizantes dirigidos contra el SRAS-CoV después de una inmunización génica con la ayuda de ADN plasmídico que dirige la expresión del ADNc de la S del SRAS-CoV.

Ejemplo 13: Aplicaciones de diagnósticos de la proteína S

La reactividad en ELISA del polipéptido recombinante Ssol se analizó frente a sueros de pacientes que padecen SRAS.

Los sueros de posibles casos de SRAS ensayados se seleccionaron sobre la base de unos resultados (positivos o negativos) de análisis de su reactividad específica frente a antígenos nativos del SRAS-CoV mediante ensayo de inmunofluorescencia sobre células VeroE6 infectadas por el SRAS-CoV y/o mediante ensayo ELISA indirecto utilizando como antígeno un lisado de células VeroE6 infectadas por el SRAS-CoV. Los sueros de estos pacientes son identificados mediante un número de orden del Centre National de Référence des virus influenzae, así como por las iniciales del paciente y el número de días transcurridos desde el principio de los síntomas. Todos los sueros de posibles casos (véase la tabla XI) reconocen los antígenos nativos del SRAS-CoV, con la excepción del suero 032552 del paciente VTT, para el cual la infección por el SRAS-CoV no se ha podido confirmar mediante RT-PCR realizada sobre extracciones respiratorias de los días 3, 8 y 12. Un panel de sueros control se usó a título de control (sueros TV): se trata de suero extraídos en Francia antes de la epidemia de SRAS aparecida en 2003.

Tabla XII: sueros de posibles casos de SRAS

suero	paciente	día de extracción
031724	JYK	7
033168	JYK	38
033597	JYK	74
032632	NTM	17
032634	THA	15
032541	PHV	10
032542	NIH	17
032552	VTT	8
032633	PTU	16
032791	JLB	3
033258	JLB	27
032703	JCM	8
033153	JCM	29

Unas fases sólidas sensibilizadas por el polipéptido recombinante Ssol se preparan mediante adsorción de una solución de polipéptido Ssol purificado a 2 µg/ml en PBS en los pocillos de una placa ELISA, después se incuban las placas durante una noche a 4°C, y se lavan con tampón PBS-Tween (PBS, 0,1% Tween20). Después de la

saturación de las placas ELISA mediante una solución de PBS-leche desnatada al 10% (peso/volumen) y del lavado en PBS-tween, los sueros a ensayar (100 µl) se diluyen al 1/400 en un tampón PBS-leche desnatada-Tween (PBS, 3% leche desnatada, 0,1% Tween), y después se añaden en los pocillos de la placa ELISA sensibilizada. Las placas se incuban durante 1h a 37°C. Después de 3 lavados con tampón PBS-Tween, se añade el conjugado anti-IgG humano marcado con la peroxidasa (ref. NA933V, Amersham) diluido al 1/4000 en tampón PBS-leche desnatada-Tween, y después las placas se incuban durante una hora a 37°C. Después de 6 lavados con tampón PBS-Tween, se añaden el cromógeno (TMB) y el sustrato (H₂O₂) y se incuban las placas durante 10 minutos protegidas de la luz. La reacción se detiene mediante la adición de una solución 1N de H₃PO₄, y después se mide la absorbencia a 450 nm, con una referencia a 620nm.

Los ensayos ELISA (figura 30) demuestran que el polipéptido recombinante Ssol está reconocido específicamente por los anticuerpos séricos de pacientes que padecen SRAS extraídos en fase media o tardía de la infección (≥ 10 días después del principio de los síntomas, mientras que no está reconocido significativamente por los anticuerpos séricos de dos pacientes (JLB y JCM) extraídos en fase precoz de la infección (3 a 8 días después del principio de los síntomas) ni por unos sueros control de sujetos que no padecen SRAS. Los anticuerpos séricos de los pacientes JLB y JCM muestran una seroconversión entre los días 3 y 27 para el primero y 8 y 29 para el segundo, después del principio de los síntomas, lo que confirma la especificidad de la reactividad de estos sueros frente al polipéptido Ssol.

En conclusión, estos resultados demuestran que el polipéptido recombinante Ssol se puede utilizar como antígeno para el desarrollo de un ensayo ELISA de diagnóstico serológico de la infección por el SRAS-CoV.

Ejemplo 14: Aplicaciones como vacuna de la proteína S soluble recombinante

Se ha estudiado la inmunogenicidad del polipéptido recombinante Ssol en ratones.

Para ello, se ha inmunizado un grupo de 6 ratones con 3 semanas de intervalo con 10 µg de polipéptido recombinante Ssol adicionado de 1 mg de hidróxido de aluminio (Alu-gel-S, Serva) diluido en PBS. Se han realizado tres inmunizaciones sucesivas y los sueros inmunes se extrajeron 3 semanas después de cada una de las inmunizaciones (IS1, IS2, IS3). A título de control, un grupo de ratones (grupo mock) ha recibido el hidróxido de aluminio sólo según el mismo protocolo.

Los sueros inmunes se ha analizado por grupos para cada uno de los 2 grupos mediante ELISA indirecto utilizando un lisado de células VeroE6 infectadas por el SRAS-CoV como antígeno, y a título de control un lisado de células VeroE6 no infectadas. Los títulos en anticuerpo anti-SRAS-CoV se calculan como la inversa de la dilución que produce una DO específica de 0,5 después del revelado por un anticuerpo policlonal anti-IgG(H+L) de ratón acoplado con la peroxidasa (NA931V, Amersham) y de TMB adicionado de H₂O₂ (KPL). Este análisis (figura 31) muestra que la inmunización por el polipéptido Ssol induce en el ratón, a partir de la primera inmunización, unos anticuerpos dirigidos contra la forma nativa de la proteína de espícula del SRAS-CoV, presente en el lisado de células VeroE6 infectadas. Después de 2 y después 3 inmunizaciones, los títulos en anticuerpos anti-S se vuelven muy elevados.

Los sueros inmunes se analizaron por grupo para cada uno de los 2 grupos para su capacidad para seroneutralizar la infectividad del SRAS-CoV. Se realizan 4 puntos de seroneutralización sobre células FRhK-4 (100 TCID₅₀ de SRAS-CoV) para cada una de las diluciones 2 veces ensayadas a partir del 1/20. El título seroneutralizante se calcula según el método de Reed y Munsch como la inversa de la dilución que neutraliza la infectividad de 2 pozos sobre 4. Este análisis muestra que los anticuerpos inducidos en el ratón por el polipéptido Ssol son neutralizantes: los títulos observados son muy elevados después de 2 y después de 3 inmunizaciones (superiores a 2560 y 5120 respectivamente, tabla XIII).

Tabla XIII: In ducción de a nticuerpos di rigidos c ontra e l SR AS-CoV de spués de la inmuni zación c on e l polipéptido recombinante Ssol. Los sueros inmunes se analizaron por grupo para cada uno de los 2 grupos para su capacidad para seroneutralizar la infectividad de 100 TCID₅₀ del SRAS-CoV sobre células FRhK-4. Se han realizado 4 puntos para cada una de las diluciones de razón 2 ensayadas a partir del 1/20. El título seroneutralizante se calcula según el método de Reed y Munsch como la inversa de la dilución que neutraliza la infectividad de 2 pozos sobre 4.

Grupo	sueros	Ac. neutralizantes
Mock	pi	< 20
	IS1	< 20
	IS2	< 20
	IS3	< 20
Ssol	pi	< 20
	IS1	57
	IS2	> 2560
	IS3	> 5120

Los títulos neutralizantes observados en los ratones inmunizados con el polipéptido Scol alcanzan niveles muy superiores a los títulos observados por Yang y otros en el ratón (2004, Nature, 428: 561-564) y a los observados por Buchholz en el hámster (2004, PNAS 101: 9804-9809), que protegen respectivamente el ratón y el hámster de la infección por el SRAS-CoV. Es por lo tanto probable que los anticuerpos neutralizantes inducidos en el ratón después de la inmunización por el polipéptido Ssol protegen estos animales contra la infección por el SRAS-CoV.

Ejemplo 15: Gen sintético optimizado para la expresión en células de mamíferos de la proteína de espícula (S) del coronavirus asociado al SRAS (SRAS-CoV).

1) Concepción del gen sintético

Se ha concebido un gen sintético que codifica la proteína de espícula del SRAS-CoV a partir del gen del aislado 031559 (plásmido pSRAS-S, C.N.C.M. n° 1-3059) a fin de permitir unos niveles de expresión elevados en unas células de mamíferos, y en particular en las células de origen humana.

Para ello:

- el uso de los codones del gen salvaje del aislado 031589 se modificó a fin de acercarse a la tendencia observada en el ser humano, y mejorar la eficacia de la traducción del ARNm correspondiente
- el contenido global en GC del gen se ha aumentado a fin de prolongar la vida media del ARNm correspondiente
- las unidades, eventualmente crípticas, susceptibles de interferir con una expresión eficaz del gen, se han suprimido (sitios donantes y receptores de corte y empalme, señales de poliadenilación, secuencias muy ricas (>80%) o muy pobres (<30%) en GC, secuencias repetidas, secuencias implicadas en la formación de estructuras secundarias del ARN, cajas TATA)
- un segundo codón STOP se ha añadido para permitir una terminación eficaz de la traducción.

Además, se han introducido unidades CpG en el gen a fin de aumentar su inmunogenecidad como vacuna de ADN. A fin de facilitar la manipulación del gen sintético, se ha colocado dos sitios de restricción BamH1 y Xho1 a ambos lados de la fase de lectura abierta de la proteína S, y los sitios de restricción BamH1, Xho1, Nhe1, Kpn1, BspE1 y Sall han sido evitados en el gen sintético.

La secuencia del gen sintético concebido (gen 040530) se da en SEC ID n° 140.

Una alineación del gen sintético 040530 con la secuencia del gen salvaje del aislado 031589 del SRAS-CoV depositada en la C.N.C.M. bajo el número 1-3059 (SEC ID n° 4, plásmido pSRAS-S) está presentado en la figura 32.

2) Construcciones plasmídicas

El gen sintético SEC ID n° 140 se ensambló a partir de oligonucleótidos sintéticos, y clonado entre los sitios Kpn1 y Sac1 del plásmido pUC-Kana para dar el plásmido 040530pUC-kana. La secuencia nucleotídica del inserto del plásmido 040530pUC-kana se verificó mediante secuenciación automática (Applied).

Se extirpó un fragmento Kpn1-Xho1 que contiene el gen sintético 040530 del plásmido 040530pUC-kana y sub-clon entre los sitios Nhe1 y Xho1 del plásmido de expresión pCI (Promega) para obtener el plásmido pCI-SSYNTH, depositado en la CNCM el 1 de diciembre de 2004, bajo el número I-3333.

Un gen sintético que codifica para la forma soluble de la proteína S, se obtuvo después fusionando las secuencias sintéticas que codifican para el ectodominio de la proteína S (aminoácidos 1 a 1193) a las de la etiqueta (FLAG: DYKDDDDK) a través de un enlazador BspE1 que codifica el dipéptido SG. Prácticamente, se ha amplificado un fragmento de ADN que codifica el ectodominio de la S del SRAS-CoV por PCR con la ayuda de los oligonucleótidos 5'-ACTAGCTAGC GGATCCACCA TGTTTCATCTT CTG-3' y 5'-AGTATCCGGAC TTGATGTACT GCTCGTACTT GC-3' a partir del plásmido 040530pUC-kana, digerido por Nhe1 y BspE1 y después insertado entre los sitios únicos Nhe1 y BspE1 del plásmido pCI-Ssol, para dar el plásmido pCI-SCUBE, depositado en la CNCM el 1 de diciembre de 2004, bajo el número I-3332. (Los plásmidos pCI-Ssol, pCI-Ssol-CTE y pCI-Ssol-WPRE (depositados en la CNCM, el 22 de noviembre de 2004, bajo el número I-3324) se obtuvieron anteriormente mediante sub-clonación del fragmento Kpn1-Xho1 extirpado del plásmido pcDNA-Ssol (véase la nota técnica de la DI 2004-106) entre los sitios Nhe1 y Xho1 de los plásmidos pCI, pCI-S-CTE y pCI-S-WPRE, respectivamente.)

Los plásmidos pCI-Scube y pCI-Ssol codifican para el mismo polipéptido recombinante Ssol.

3) Resultados

La capacidad del gen sintético que codifica la proteína S para dirigir eficazmente la expresión de la S del SRAS-CoV en unas células de mamíferos se ha comparado con la del gen salvaje después de la transfección transitoria de las células de primates (VeroE6), y de células humanas (293T).

En el experimento representado por la figura 33, y en la tabla XIV, se han transfectado unas monocapas de 5×10^5 células VeroE6 o 7×10^5 células 293T en cajas de Petri de 35 mm con 2 μ g de plásmidos pCI (a título de control), pCI-S, pCI-S-CTE, pCI-S-WPRE y pCI-Ssynth y 6 μ l de agente reactivo Eugene6 según las indicaciones del fabricante (Roche). Después de 48 horas de incubación a 37°C y bajo 5% de CO₂, se han preparado unos extractos celulares en tampón de depósito según Laemmli, separados sobre un gel SDS al 8% de poliácridamida y después transferidos sobre una membrana de PVDF (BioRad). La detección de esta inmunohuella («western blot») se ha realizado con la ayuda de un suero policlonal de conejo anti-S (suero inmune del conejo P11135: véase el Ejemplo 4 anterior) y de anticuerpos policlonales de asno dirigidos contra las IgG de conejo y acoplados con la peroxidasa (NA934V, Amersham). La inmunohuella se revela de manera cuantitativa por luminiscencia con la ayuda del kit ECL+ (Amersham), y adquisición sobre un dispositivo de imágenes digitales (FluorS, BioRad).

El análisis de los resultados obtenidos con el programa QuantityOne v4.2.3 (BioRad) muestra que en este experimento, el plásmido pCI-Synth permite la expresión transitoria de la proteína S a niveles elevados en las células VeroE6 y 293T, mientras que el plásmido pCI-S no permite inducir una expresión a niveles suficientes para ser detectada. Los niveles de expresión observados son del orden de 2 veces superiores a los observados con el plásmido pCI-S-WPRE.

Tabla XIV: Utilización de un gen sintético para la expresión de la S del SRAS-CoV. Se separaron unos extractos celulares preparados 48 horas después de la transfección de células VeroE6 o 293T por los plásmidos pCI, pCI-S, pCI-S-CTE, pCI-S-WPRE y pCI-Ssynth sobre un gel SDS al 8% de acrilamida y analizados mediante transferencia western con la ayuda de un anticuerpo policlonal de conejo anti-S y de un anticuerpo policlonal anti-IgG(H+L) de conejo acoplado con la peroxidasa (NA934V, Amersham). La transferencia western se revela mediante luminiscencia (ECL+, Amersham), y la adquisición sobre un dispositivo de imágenes digitales (FluorS, BioRad). Los niveles de expresión de la proteína S se han medido cuantificando las 2 bandas mayoritarias localizadas en la imagen (véase la figura 33), y son indicados en función de una escala arbitraria en la que el valor de 1 representa el nivel medido después de la transfección del plásmido pCI-S-WPRE.

plásmido	VeroE6	293T
pCI	0,0	0,0
pCI-S	< 0,1	< 0,1
pCI-S-CTE	0,5	< 0,1
pCI-S-WPRE	1,0	1,0
pCI-Ssynth	1,8	1,9

En una segunda etapa, la capacidad del gen sintético Scube para dirigir eficazmente la síntesis y la secreción del polipéptido Ssol por unas células de mamíferos se ha comparado a la del gen salvaje después de la transfección transitoria de células de hámster (BHK-21) y de células humanas (293T).

En el experimento presentado por la tabla XV, se han transfectado unas monocapas de 6×10^5 células BHK-21 y de 7×10^5 células 293T en cajas de Petri de 35 mm con 2 μ g de los plásmidos pCI (a título de control), pCI-Ssol, pCI-Ssol-CTE, pCI-Ssol-WPRE y pCI-Scube y 6 μ l de agente reactivo Eugene6 según las indicaciones del fabricante (Roche). Después de 48 horas de incubación a 37°C y bajo 5% de CO₂, los sobrenadantes celulares se extrajeron y analizaron de manera cuantitativa para la secreción del polipéptido Ssol mediante un ensayo ELISA-captura específico del polipéptido Ssol.

El análisis de los resultados muestra que, en este experimento, el plásmido pCI-Scube permite la expresión del polipéptido Ssol a niveles 8 veces (células BHK-21) a 20 veces (células 293T) más elevados que el plásmido pCI-Ssol. Los niveles de expresión observados son del orden de 2 veces (células 293T) a 5 veces (células BHK-21) superiores a los observados con el plásmido pCI-Ssol-WPRE.

Tabla XV: Utilización de un gen sintético para la expresión del polipéptido Ssol. Los sobrenadantes se recogieron 48 horas después de la transfección de células BHK o 293T por los plásmidos pCI, pCI-Ssol, pCI-Ssol-CTE, pCI-Ssol-WPRE y pCI-Scube y analizados de manera cuantitativa para la secreción del polipéptido Ssol mediante un ensayo ELISA-captura específico del polipéptido Ssol. Las transfecciones se han realizado en duplicado y los resultados se indican en forma de medias y desviación típica de las concentraciones de polipéptido Ssol (ng/ml) medidas en los sobrenadantes.

Plásmido	BHK	293T
Pci	< 20	< 20
pCI-Ssol	< 20	56 ± 10
pCI-Ssol-CTE	< 20	63 ± 8
pCI-Ssol-WPRE	28 ± 1	531 ± 15
PCI-Scube	152 ± 6	1140 ± 20

En resumen, estos resultados muestran que la expresión en unas células de mamíferos del gen sintético 040530 que codifica la S del SRAS-CoV bajo la dependencia de secuencias promotoras del ARN polimerasa II es mucho más eficaz que la del gen salvaje del aislado 031589. Esta expresión es incluso más eficaz que la dirigida por el gen salvaje en presencia de las secuencias WPRE del virus de la hepatitis de la marmota.

4) Aplicaciones

El uso del gen sintético 040530 que codifica la S del SRAS-CoV, o de su variante Scube que codifica el polipéptido Ssol es susceptible de sustituir ventajosamente el gen salvaje en numerosas aplicaciones, en las que la expresión de la S es necesaria a niveles elevados. En particular, para:

- mejorar la eficacia de la inmunización génica por unos plásmidos del tipo pCI-Ssynth incluso pCI-Ssynth-CTE o pCI-Ssynth-WPRE
- establecer nuevas líneas celulares que expresan cantidades más elevadas de la proteína S o del polipéptido Ssol con la ayuda de vectores lentivirales recombinantes portadores del gen Ssynth o del gen Scube respectivamente
- mejorar la inmunogenicidad de los vectores lentivirales recombinantes que permite la expresión de la proteína S o del polipéptido Ssol
- mejorar la inmunogenicidad de vectores vivos que permiten la expresión de la proteína S o del polipéptido Ssol como unos virus recombinantes de la vacuna antivariólica, o unos virus de la rubeola recombinantes (véanse los ejemplos ejemplos 16 y 17 a continuación)

Ejemplo 16: Expresión de la proteína de espícula (S) del coronavirus asociado al SRAS (SRAS-CoV) con la ayuda de virus recombinantes de la vacuna antivariólica.

Aplicación vacuna.

Aplicación a la producción de una forma soluble de la proteína de espícula (S) y concepción de un ensayo serológico del SRAS.

1) Introducción

El objetivo de este ejemplo es evaluar la capacidad de virus recombinantes de la vacuna antivariólica (VV), que expresa diferentes antígenos del coronavirus asociado al SRAS (SRAS-CoV) para constituir nuevos candidatos vacvíneos contra el SRAS y un medio para producir unos antígenos recombinantes en células de mamíferos.

Para ello, los inventores se han interesado por la proteína de espícula (S) del SRAS-CoV, que permite inducir después de la inmunización génica en el animal, unos anticuerpos que neutralizan la infectividad del SRAS-CoV, así como a una forma soluble y secretada de esta proteína, el polipéptido Ssol, que es un compuesto del ectodominio (aa 1-1193) de la S fusionado en su extremo C-ter a una etiqueta FLAG (DYKDDDDK) vía un enlazador BspE1 que codifica el dipéptido SG. Este polipéptido Ssol presenta una antigenicidad similar a la de la proteína S y permite, después de la inyección al ratón en forma de una proteína purificada adicionada en hidróxido de aluminio, la inducción de títulos elevados de anticuerpos neutralizantes contra el SRAS-CoV.

Las diferentes formas del gen S se han colocado bajo la dependencia del promotor del gen 7.5K y después se han introducido en el locus de la timidina kinasa (TK) de la cepa Copenhague del virus de la vacuna antivariólica mediante doble recombinación homóloga *in vivo*. A fin de mejorar la inmunogenicidad de los virus de la vacuna antivariólica recombinantes, se ha seleccionado un promotor tardío sintético en lugar del promotor 7.5K, para aumentar la producción de S y Ssol durante unas fases tardías del ciclo viral.

Después de aislar los virus de la vacuna antivariólica recombinantes y de verificar su capacidad para expresar el antígeno S del SRAS-CoV, se ha ensayado su capacidad para inducir en el ratón una respuesta inmunitaria contra el SRAS. Después de haber purificado el antígeno Ssol del sobrenadante de células infectadas, se ha concebido un

ensayo ELISA de serodiagnóstico del SRAS, y se ha evaluado su eficacia con la ayuda de sueros de posibles casos de SRAS.

2) Construcción de virus recombinantes

Se obtuvieron virus recombinantes de la vacuna antivariólica, que dirigen la expresión de la glicoproteína S del aislado 031589 del SRAS-CoV y de forma soluble y secretada de esta proteína, el polipéptido Ssol, bajo la dependencia del promotor 7.5K. Con el objetivo de aumentar los niveles de expresión de S y Ssol, se obtuvieron también unos virus recombinantes, en los que los ADNc de S y de Ssol están dispuestos bajo la dependencia de un promotor sintético tardío.

El plásmido pTG186poli es un plásmido de transferencia para la construcción de virus recombinantes de la vacuna antivariólica (Kieny, 1986, Biotechnology, 4:790-795). Para ello, contiene el gen de la timidina kinasa del VV en el que se ha insertado el promotor del gen 7.5K, seguido de un sitio múltiple de clonación que permite la inserción de genes heterólogos (figura 34A). El promotor del gen 7.5K contiene un grupo de dos secuencias promotoras activas respectivamente durante las fases precoces (P_E) y tardías (P_L) del ciclo de replicación del virus de la vacuna antivariólica. Los fragmentos BamH1-Xho1 se han extirpado de los plásmidos pTRIP-S y pcDNA-Ssol respectivamente e insertados entre los sitios BamH1 y Sma1 del plásmido pTG186poli para dar los plásmidos pTG-S y pTG-Ssol (figura 34A). Los plásmidos pTG-S y pTG-Ssol se han depositado en la CNCM, el 2 de diciembre de 2004, bajo los números I-3338 e I-3339, respectivamente.

Los plásmidos pTN480, pTN-S y pTN-Ssol se obtuvieron a partir de los plásmidos pTG186poli, pTG-S y pTG-Ssol respectivamente, sustituyendo el fragmento Nde1-Pst1 que contiene el promotor 7.5K por un fragmento de ADN que contiene el promotor tardío sintético 480, que se obtuvo mediante hibridación de los oligonucleótidos 5'-TATGAGCTTT TTTTTTTTTT TTTTTTGGC ATATAAATAG ACTCGGCGCG CCATCTGCA-3' y 5'-GATGGCGCGC CGAGTCTATT TATATGCCAA AAAAAAAAAA AAAAAAAGC TCA-3' (figura 34B). El inserto se ha secuenciado con la ayuda de un kit BigDye Terminator v1.1 (Applied Biosystems) y de un secuenciador automático ABI377. La secuencia del promotor sintético tardío 480 tal como se clonó en los plásmidos de transferencia de la serie pTN está indicada en la figura 34C. Los plásmidos pTN-S y pTN-Ssol se han depositado en la CNCM, el 2 de diciembre de 2004, bajo los números I-3340 e I-3341, respectivamente.

Los virus recombinantes de la vacuna antivariólica se obtuvieron mediante doble recombinación homóloga *in vivo* entre el casete TK de los plásmidos de transferencia de las series pTG y pTN y el gen TK de la cepa Copenhagen del virus de la vacuna antivariólica según un procedimiento descrito por Kieny y otros (1984, Nature, 312: 163-166). Brevemente, se transfectan unas células CV-1 con la ayuda de DOTAP (Roche) mediante el ADN genómico de la cepa Copenhagen del virus de la vacuna antivariólica y cada uno de los plásmidos de transferencia de las series pTG y pTN descritos anteriormente, y después se sobreinfecta por el virus de la vacuna antivariólica auxiliar VV-ts7 durante 24 horas a 33°C. El virus auxiliar está contra-seleccionado por incubación a 40°C durante 2 días, y después los virus recombinantes (fenotipo TK-) se seleccionan mediante dos ciclos de clonación en medio agar sobre células 143Btk- en presencia de BuDr (25 µg/ml). Los 6 virus VV-TG, VV-TG-S, VV-TG-Ssol, VV-TN, VV-TN-S y VV-TN-Ssol se obtuvieron respectivamente con la ayuda de los plásmidos de transferencia pTG186poli, pTG-S, pTG-Ssol, pTN480, pTN-S, pTN-Ssol. Los virus VV-TG y VV-TN no expresan ningún gen heterólogo y han sido utilizados como control TK- en los experimentos. Las preparaciones de virus recombinantes se han realizado sobre monocapas de células CV-1 o BHK-21, y el título en unidades que forma el intervalo (u.f.p.) determinado sobre células CV-1 según Earl y Moss (1998, Current Protocols in Molecular Biology, 16.16.1-16.16.13).

3) Caracterización de los virus recombinantes

Se ha buscado la expresión de los transgenes que codifican la proteína S y el polipéptido Ssol mediante transferencia western.

Se han infectado monocapas de células CV-1 con una multiplicidad de 2 por los diferentes virus de la vacuna antivariólica recombinantes VV-TG, VV-TG-S, VV-TG-Ssol, VV-TN, VV-TN-S y VV-TN-Ssol. Después de 18 horas de incubación a 37°C y bajo 5% de CO₂, se han preparado unos extractos celulares en tampón de depósito según Laemmli, se han separado sobre gel SDS al 8% de poli(acrilamida), y después se han transferido sobre una membrana de PVDF (BioRad). La detección de esta inmunohuella («western blot») se ha realizado con la ayuda de un suero policlonal de conejo anti-S (suero inmune del conejo P11135: véase el ejemplo 4) y de anticuerpos policlonales de asno dirigidos contra las IgG de conejo, y acoplados con la peroxidasa (NA934V, Amersham). Los anticuerpos fijados se han revelado mediante luminescencia con la ayuda del kit ECL+ (Amersham) y de películas de autorradiografía Hyperfilm MP (Amersham).

Como se muestra en la figura 35A, el virus recombinante VV-TN-S dirige la expresión de la proteína S a niveles que son comparables a los que se puede observar 8h después de la infección por el SRAS-CoV pero que son mucho más elevados que los que se pueden observar después de la infección por VV-TG-S. En un segundo experimento (figura 35B), el análisis de cantidades variables de extractos celulares muestra que los niveles de expresión observados después de la infección por los virus de la serie TN (VV-TN-S y VV-TN-Ssol) son aproximadamente 10

veces más elevados que los observados con los virus de la serie TG (VV-TG-S y W-TG-Ssol respectivamente). Además, el polipéptido Ssol está secretado en el sobrenadante de células CV-1 infectadas por el virus VV-TN-Ssol más eficazmente que en el sobrenadante de células infectadas por VV-TG-Ssol (figura 36A). En este experimento, el virus VV-TN-Sflag se utilizó a título de control, porque expresa la forma membranaria de la proteína S fusionada en su extremo C-ter a la etiqueta FLAG. La proteína Sflag no está detectada en el sobrenadante de las células infectadas por VV-TN-Sflag, lo que demuestra que el polipéptido Ssol está bien secretado de manera activa después de la infección por W-TN-Ssol.

Estos resultados demuestran que los virus de la vacuna antivariólica recombinantes son portadores de los transgenes, y permiten la expresión de la glicoproteína del SRAS en su forma membranaria (S) o en una forma soluble y secretada (Ssol). El virus de la vacuna antivariólica portador del promotor sintético 480 permite la expresión de S y la secreción de Ssol a niveles mucho más elevados que los virus portadores del promotor del gen 7.5K.

4) Aplicación a la producción de una forma soluble de la S del SRAS-CoV.

Purificación de este antígeno recombinante y aplicaciones de diagnósticos.

La línea BHK-21 es la línea celular que segregan las cantidades más elevadas de polipéptido Ssol después de la infección por el virus VV-TN-Ssol entre las líneas ensayadas (BHK-21, CV1, 293T y FrhK-4, figura 36B); permite la producción y la purificación en cantidad del polipéptido recombinante Ssol. En un experimento del tipo en el que las condiciones experimentales de infección, de producción y de purificación se han optimizado, las células BHK-21 están inoculadas en un medio de cultivo estándar (DMEM sin piruvato que contiene 4,5 g/l de glucosa y suplementado por 5% de TPB, 5% de SVF, 100 U/ml de penicilina y 100 µg/ml de estreptomycin) en forma de una monocapa sub-confluente (10 millones de células para cada 100 cm² en 25 ml de medio). Después de 24h de incubación a 37°C bajo 5% de CO₂, las células son infectadas a una M.O.I. de 0,03 y el medio estándar se sustituye por el medio de secreción en el que la cantidad de SVF está disminuido al 0,5% y el TPB está suprimido. El sobrenadante de cultivo se extrae después de 2,5 días de incubación a 35°C y bajo 5% de CO₂ y el virus de la vacuna antivariólica se inactiva mediante adición de tritón X-100 (0,1%). Después de la filtración sobre membrana de polietersulfona (PES) de 0,1 µm, el polipéptido recombinante Ssol se purifica mediante una cromatografía de afinidad sobre la plantilla anti-FLAG con elución por una solución de péptido FLAG (DYKDDDDK) a 100 µg/ml en TBS (Tris 50mM pH 7,4, 150 mMNaCl).

El análisis por gel SDS al 8% de acrilamida coloreado con nitrato de plata ha puesto en evidencia un polipéptido mayoritario cuya masa molecular es de aproximadamente 180kD y cuyo grado de pureza es superior al 90% (figura 37). La concentración del polipéptido recombinante Ssol purificado se determinó mediante comparación con los marcadores de masa molecular, y se estima a 24 ng/µl.

Esta preparación de polipéptido Ssol purificado permite realizar una escala estándar para medir, con la ayuda de un ensayo ELISA-captura, las concentraciones de Ssol presentes en los sobrenadantes de cultivo. Según este ensayo, la línea BHK-21 segrega aproximadamente 1 µg/ml de polipéptido Ssol en condiciones de producción descritas anteriormente. Además, el esquema de purificación presentado permite purificar del orden de 160 µg de polipéptido Ssol por litro de sobrenadante de cultivo.

La reactividad en ELISA del polipéptido recombinante Ssol se ha analizado frente a sueros de pacientes que padecen STRAS.

Los sueros de posibles casos de SRAS ensayados se seleccionaron en base a resultados (positivos o negativos) de análisis de su reactividad específica frente a antígenos nativos del SRAS-CoV mediante ensayo de inmunofluorescencia sobre células VeroE6 infectadas por el SRAS-CoV y/o mediante el ensayo ELISA indirecto utilizando como antígeno un lisado de células VeroE6 infectadas por el SRAS-CoV. Los sueros de estos pacientes son identificados por un número de orden del Centre National de Référence des virus influenzae, así como mediante las iniciales del paciente, y el número de días transcurridos desde el principio de los síntomas. Todos los sueros de posibles casos (véase la tabla XVI) reconocen los antígenos nativos del SRAS-CoV con la excepción del suero 032552 del paciente VTT, para el cual la infección por el SRAS-CoV no se ha podido confirmar por RT-PCR realizada sobre extracciones respiratorias de los días 3, 8 y 12. Un panel de sueros control se ha utilizado a título de control (sueros TV): se trata de sueros extraídos en Francia antes de la epidemia de SRAS aparecida en 2003.

Tabla XVI: sueros de posibles casos de SRAS

suero	paciente	Día de extracción
033168	JYK	38
033597	JYK	74
032632	NTM	17

032634	THA	15
032541	PHV	10
032542	NIH	17
032552	VTT	8
032633	PTU	16

Se han preparado unas fases sólidas sensibilizadas por el polipéptido recombinante Ssol mediante adsorción de una solución de polipéptido Ssol purificado a 4 µg/ml en PBS en los pocillos de una placa ELISA. Las placas se incuban durante una noche a 4°C y después se lavan con tampón PBS-Tween (PBS, 0,1% Tween20). Después del lavado en PBS-tween, los sueros a ensayar (100 µl) se diluyen al 1/100 y al 1/400 en tampón PBS-leche desnatada-Tween (PBS, 3% leche desnatada, 0,1% Tween) y después se añaden en los pocillos de la placa ELISA sensibilizada. Las placas se incuban durante 1h a 37°C. Después de 3 lavados con tampón PBS-Tween, se añade el conjugado anti-IgG humano marcado con la peroxidasa (ref NA933V, Amersham) diluido al 1/4000 en tampón PBS-leche desnatada-Tween, y después se incuban las placas una hora a 37°C. Después de 6 lavados con tampón PBS-Tween, se añaden el cromógeno (TMB) y el sustrato (H₂O₂), y se incuban las placas durante 10 minutos protegidas de la luz. La reacción se detiene mediante adición de una solución 1M de H₃PO₄, y después se mide la absorbencia a 450 nm con una referencia a 620 nm.

Los ensayos ELISA (figura 38) demuestran que el polipéptido recombinante Ssol está reconocido específicamente por los anticuerpos séricos de pacientes que padecen SRAS extraídos en fase media o tardía de la infección (≥ 10 días después del principio de los síntomas), mientras que no se ha reconocido de manera significativa por los anticuerpos séricos de los sueros controles de sujetos que no padecen SRAS.

En conclusión, estos resultados demuestran que el polipéptido recombinante Ssol puede ser purificado a partir del sobrenadante de células de mamíferos infectadas por el virus de la vacuna antivariólica recombinante VV-TN-Ssol y ser utilizado como antígeno para la realización de un ensayo ELISA de diagnóstico serológico de la infección por el SRAS-CoV.

5) Aplicaciones vacunas

Se ha estudiado la inmunogenicidad de los virus de la vacuna antivariólica recombinantes en el ratón.

Para ello, se han inmunizado unos grupos de 7 ratones BALB/c por vía i.v. dos veces con 4 semanas de intervalo por 10⁶ u.f.p. de virus de la vacuna antivariólica recombinantes VV-TG, VV-TG-S, VV-TG-Ssol, VV-TN, VV-TN-S y VV-TN-Ssol así como, a título de control, VV-TG-HA que dirige la expresión de la hemaglutinina de la cepa A/PR/8/34 del virus de la gripe. Los sueros inmunes se han extraído 3 semanas después de cada una de las inmunizaciones (IS1, IS2).

Los sueros inmunes se han analizado por grupos para cada uno de los grupos por ELISA indirecto, utilizando un lisado de células VeroE6 infectadas por el SRAS-CoV como antígeno y a título de control, un lisado de células VeroE6 no infectadas. Los títulos (TI) en anticuerpos anti-SRAS-CoV se calculan como la inversa de la dilución que produce una DO específica de 0,5 después del revelado por un anticuerpo policlonal anti-IgG(H+L) de ratón acoplado con la peroxidasa (NA931 V, Amersham), y de TMB adicionado de H₂O₂ (KPL). Este análisis (figura 39A) muestra que la inmunización por el virus VV-TG-S y VV-TN-S induce en el ratón, a partir de la primera inmunización, unos anticuerpos dirigidos contra la forma nativa de la proteína de espícula del SRAS-CoV presente en el lisado de células VeroE6 infectadas. Las respuestas inducidas por el virus VV-TN-S son más elevadas que las inducidas por el virus VV-TG-S después de la primera (TI=740 y TI=270 respectivamente) y de la segunda (TI=3230 y TI=600 respectivamente) inmunización. El virus VV-TN-Ssol induce fuertes títulos de anticuerpos anti-SRAS-CoV después de dos inmunizaciones (TI=640), mientras que el virus VV-TG-Ssol induce una respuesta al límite de la detección (TI=40).

Se han analizado los sueros inmunes por grupos para cada uno de los grupos para su capacidad a seroneutralizar la infectividad del SRAS-CoV. Se realizan 4 puntos de seroneutralización sobre células FRhK-4 (100 TCID₅₀ de SRAS-CoV) para cada una de las diluciones 2 veces ensayadas a partir del 1/20. El título seroneutralizante se calcula según el método de Reed y Munsch como la inversa de la dilución que neutraliza la infectividad de 2 pozos sobre 4. Este análisis muestra que los anticuerpos inducidos en el ratón por los virus de la vacuna antivariólica que expresa la proteína S o el polipéptido Ssol son netralizantes y que los virus de promotores sintéticos son unos inmunógenos más eficaces que los virus portadores del promotor 7.5K: los títulos más elevados (640) son observados después de 2 inmunizaciones por el virus VV-TN-S (figura 39B).

El poder protector de los anticuerpos neutralizantes inducidos en el ratón después de la inmunización por los virus de la vacuna antivariólica recombinantes se evalúa con la ayuda de una infección de prueba por el SRAS-CoV.

6) Otras aplicaciones

Se construyen virus de la vacuna antivariólica recombinantes de tercera generación sustituyendo las secuencias salvajes de los genes S y Ssol por unos genes sintéticos optimizados para la expresión en células de mamíferos, descritos anteriormente. Estos virus de la vacuna antivariólica recombinantes son susceptibles de expresar unas cantidades más importantes de antígenos S y Ssol, y por lo tanto una inmunogenicidad incrementada.

El virus recombinante de la vacuna antivariólica VV-TN-Ssol se puede utilizar para la producción en cantidad y la purificación del antígeno Ssol, para aplicaciones de diagnósticos (serología por ELISA) y vacunas (vacuna sub-unitaria).

Ejemplo 1.7: Virus recombinante de la rubeola que expresa la proteína de espícula (S) del coronavirus asociado al SRAS (SRAS-CoV). Aplicaciones vacunas.

1) Introducción

La vacuna de la rubeola (MV) induce en el ser humano una inmunidad protectora de larga duración después de una sola inyección (Hilleman, 2002, Vaccine, 20: 651-665). La protección conferida es muy firme y se basa en la inducción de una respuesta en anticuerpos y de una respuesta celular CD4 y CD8. El genoma del MV es muy estable y no se ha observado jamás ninguna reversión hacia la virulencia de las cepas de vacuna. El virus de la rubeola pertenece al género de los *Morbillivirus* de la familia de los *Paramyxoviridae*; es un virus recubierto cuyo genoma es un ARN monocatenario de polaridad negativa de 16kb (figura 40A) y cuyo ciclo de replicación exclusivamente citoplásmico excluye cualquier posibilidad de integración en el genoma del hospedante. La vacuna de la rubeola es así una de las vacunas vivas más eficaz y más segura utilizadas en la población humana. El equipo de Frédéric Tangy ha desarrollado recientemente un vector de expresión en base a la cepa Schwarz del virus de la rubeola, que es la cepa atenuada más segura y más utilizada en el ser humano como vacuna contra la rubeola. Esta cepa vacuna puede ser aislada a partir de un clon molecular infeccioso, conservando al mismo tiempo su inmunogenicidad en los primates así como el ratón susceptibles a la infección. Constituye, después de la inserción de unidades de transcripción suplementarias, un vector para la expresión de secuencias heterólogas (Combret, 2003, J. Virol. 77: 11546-11554). Además, un MV Schwarz recombinante que expresa la glicoproteína de envoltura del virus West Nile (WNV) induce una respuesta en anticuerpos eficaz y de larga duración que protege el ratón de una infección de prueba letal por el WNV (Despres y otros, 2004, J. Infect. Dis., en impresión). Todas estas características hacen de la cepa atenuada Schwarz del virus de la rubeola un candidato vector extremadamente prometedor para la construcción de nuevas vacunas vivas recombinantes.

El objetivo de este ejemplo es evaluar la capacidad de virus recombinantes de la rubeola (MV) que expresa diferentes coronavirus asociado al SRAS (SRAS-CoV) para constituir nuevos candidatos vacvíneos contra el SRAS.

Los inventores se han interesado en la proteína de espícula (S) del SRAS-CoV, que permite inducir después de la inmunización génica en el animal unos anticuerpos que neutralizan la infectividad del SRAS-CoV, así como una forma soluble y secretada de esta proteína, el polipéptido Ssol, que está compuesto del ectodominio (aa 1-1193) de la S fusionado en su extremo C-ter a una etiqueta FLAG (DYKDDDDK) a través de un enlazador BspE1 que codifica el dipéptido SG. Este polipéptido Ssol presenta una antigenicidad similar a la de la proteína S y permite, después de la inyección al ratón en forma de una proteína purificada adyuvada en hidróxido de aluminio, la inducción de títulos elevados de anticuerpos que neutralizantes contra el SRAS-CoV.

Las diferentes formas del gen S se han introducido en forma de una unidad de transcripción suplementaria entre los genes P (fosfoproteína) y M (plantilla) en el ADNc de la cepa Schwarz del MV anteriormente descrito (Combret, 2003, J. Virol. 77: 11546-11554; solicitud EP N° 02291551.6 del 20 de junio de 2002, y solicitud EP N° 02291550.8 del 20 de junio de 2002). Después de aislar los virus recombinantes MVSchw2-SRAS-S y MVSchw2-SRAS-Ssol y verificar su capacidad para expresar el antígeno S del SRAS-CoV, se ensaya su capacidad para inducir en el ratón y después en el mono una respuesta inmunitaria protectora contra el SRAS.

2) Construcción de los virus recombinantes

El plásmido pTM-MVSchw-ATU2 (figura 40B) contiene un ADNc infeccioso que corresponde al antígenoma de la cepa vacuna Schwarz del virus de la rubeola (MV) en el que una unidad de transcripción suplementaria (ATU) se ha introducido entre los genes P (fosfoproteína) y M (plantilla) (Combret, 2003, Journal of Virology, 77: 11546-11554). Se han construido unos genomas recombinantes MVSchw2-SRAS-S y MVSchw2-SRAS-Ssol del virus de la rubeola mediante la inserción de los ORF de la proteína S, y del polipéptido Ssol en la unidad de transcripción suplementaria del vector MVSchw-ATU2.

Para ello, un fragmento de ADN que contiene el ADNc de la S del SRAS-CoV se ha amplificado mediante PCR con la ayuda de los oligonucleótidos 5'-ATACGTACGA CCATGTTTAT TTTCTTATTA TTTCTTACTC TCACT-3' y 5'-ATAGCGCGCT CAT- TATGTGT AATGTAATTT GACACCTTG-3' utilizando el plásmido ADNpc-S como plantilla, y

después insertado en el plásmido pCR®2.1-TOPO (Invitrogen) para obtener el plásmido pTOPO-S-MV. Los dos oligonucleótidos utilizados contienen unos sitios de restricción BsiW1 y BssHII, a fin de permitir la inserción ulterior en el vector de la rubeola, y se han concebido a fin de generar una secuencia de 3774 nt que incluye los codones de iniciación y de terminación, a fin de respetar la regla de los 6 que estipula que la longitud del genoma de un virus de la rubeola debe ser divisible por 6 (Calain & Roux, 1993, J. Virol., 67: 4822-4830; Schneider y otros, 1997, Virology, 227: 314-322). El inserto se ha secuenciado con la ayuda de un kit BigDye Terminator v1.1 (Applied Biosystems) y de un secuenciador automático ABI377.

Con el fin de expresar una forma soluble y secretada de la S del SRAS-CoV, se ha obtenido después un plásmido que contiene el ADNc del polipéptido Ssol que corresponde al ectodominio (aa 1-1193) de la S del SRAS-CoV fusionado en su extremo C-ter a la secuencia de una etiqueta FLAG (DYKDDDDK) a través de un enlazador BspE1 que codifica el dipéptido SG. Para ello, se ha amplificado un fragmento de ADN con la ayuda de los oligonucleótidos 5'-CCATTTC AAC AATTTGGCCG-3' y 5'- ATAGGATCCG CGCGTCATT ATTTATCGTC GTCATCTTTATAATC-3' a partir del plásmido pCDNA-Ssol, y después insertado en el plásmido pTOPO-S-MV entre los sitios Sal1 y BamH1 para obtener el plásmido pTOPO-S-MV-SF. La secuencia generada es larga de 3618 nt entre los sitios BsiW1 y BssHII y respeta la regla de los 6. El inserto se ha secuenciado como se ha indicado anteriormente.

Los fragmentos BsiW1-BssHII que contienen los ADNc de la proteína S y del polipéptido Ssol se han extirpado después mediante digestión de los plásmidos pTOPO-S-MV y pTOPO-S-MV-SF y después se han sub-clonado entre los sitios correspondientes del plásmido pTM-MV Schw-ATU2 para dar los plásmidos pTM-MV Schw2-SRAS-S y pTM-MV Schw2-SRAS-Ssol (figura 40B). Estos dos plásmidos se han depositado en la C.N.C.M. el 1 de diciembre de 2004, bajo los números 1-3326 y 1-3327, respectivamente.

Los virus de la rubeola recombinantes que corresponden a los plásmidos pTM-MV Schw2-SRAS-S y pTM-MV Schw2-SRAS-Ssol se han obtenido mediante genética inversa según el sistema que se basa en el uso de una línea celular auxiliar, descrito por Radecke y otros (1995, Embo J., 14: 5773-5784) y modificado por Parks y otros (1999, J. Virol., 73: 3560-3566). Brevemente, las células auxiliares 293-3-46 son transfectadas según el método con fosfato de calcio por 5 µg de los plásmidos pTM-MV Schw2-SRAS-S o pTM-MV Schw2-SRAS-Ssol y 0,02 µg del plásmido pEMC-La que dirige la expresión de la polimerasa L del MV (donación de M. A. Billeter). Después de una noche de incubación a 37°C, se realiza un choque térmico durante 2 horas a 43°C y las células transfectadas son transferidas sobre una monocapa de células Vero. Para cada uno de los dos plásmidos, unos sincitios han aparecido después de 2 a 3 días de cocultivo, y se han transferido sucesivamente sobre una monocapas de células Vero al 70% de confluencia en cajas de petri de 35 mm y después en frascos de 25 y 75 cm². Cuando los sincitios alcanzan 80-90% de confluencia, se recuperan las células con la ayuda de un raspador, y después se congelan y se descongelan una vez. Después de una centrifugación a baja velocidad, el sobrenadante que contiene el virus se conserva alícuotamente a -80°C. Los títulos de los virus recombinantes MV Schw2-SRAS-S y MV Schw2-SRAS-Ssol se han determinado mediante dilución límite sobre células Vero y el título en dosis que infecta al 50% los pozos (TCID₅₀) se ha calculado según el método de Karber.

3) Caracterización de los virus recombinantes

Se ha buscado la expresión de los transgenes que codifican la proteína S y el polipéptido Ssol mediante transferencia western e inmunofluorescencia.

Se han infectado unas monocapas de células Vero en frascos T-25 a una multiplicidad de 0,05 mediante diferentes pasos de los dos virus MV Schw2-SRAS-S y MV Schw2-SRAS-Ssol y el virus salvaje MWSchw a título de control. Cuando los sincitios alcanzan 80 a 90% de confluencia, se preparan unos extractos citoplásmicos en un tampón de extracción (150 mM NaCl, 50 mM Tris-HCl pH 7,2, 1% tritón-X-100, 0,1% SDS, 1% DOC) y después se diluye en tampón de depósito según Laemmli, se separan sobre un gel SDS al 8% de poliácridamida y se transfieren sobre una membrana de PVDF (BioRad). La detección de esta inmunohuella («western blot») se realiza con la ayuda de un suero policlonal de conejo anti-S (suero inmune del conejo P11135: véase el ejemplo 4 anterior) y de anticuerpos policlonales de asno dirigidos contra las IgG de conejo y acoplados con la peroxidasa (NA934V, Amersham). Los anticuerpos fijados se revelan mediante luminiscencia con la ayuda del kit ECL+ (Amersham) y de películas de autorradiografía Hyperfilm MP (Amersham).

Se han infectado células Vero en monocapas sobre láminas de vidrio mediante los dos virus MV Schw2-SRAS-S y MV Schw2-SRAS-Ssol y el virus salvaje MWSchw a título de control a multiplicidades de infección de 0,05. Cuando los sincitios alcanzan del 90 al 100% (virus MV Schw2-SRAS-Ssol) o del 30 al 40% (MV Schw2-SRAS-S, MWSchw) de confluencia, las células se fijan en una solución de PBS-PFA 4%, se permeabilizan mediante una solución de PBS que contiene 0,2% de tritón y después se marcan mediante anticuerpos policlonales de conejos hiperinmunizados por medio de los viriones purificados y inactivados del SRAS-CoV, y mediante un conjugado de anticuerpos de cabra anti-IgG(H+L) de conejo acoplado al FITC (Jackson).

Como se muestra en las figuras 41 y 42, los virus recombinantes MV Schw2-SRAS-S y MV Schw2-SRAS-Ssol dirigen la expresión de la proteína S y del polipéptido Ssol respectivamente a niveles comparables a los que se pueden observar 8h después de la infección por el SRAS-CoV. La expresión de estos polipéptidos es estable después de 3

pasos de los virus recombinantes en cultivo celular. Estos resultados demuestran que los virus de la rubeola recombinantes son bien portadores de los transgenes y permiten la expresión de la glicoproteína del SRAS en su forma membranaria (S) o en una forma soluble (Ssol). Se espera que el polipéptido Ssol sea secretado de las células infectadas por el virus MVSchw2-SRAS-Ssol, tal como es el caso cuando este mismo polipéptido está expresado en unas células de mamíferos después de la transfección transitoria de las secuencias correspondientes (véase el ejemplo 11 anterior).

4) Aplicaciones

Como se ha demostrado que los virus MVSchw2-SRAS-S y MVSchw2-SRAS-Ssol permiten la expresión de la S del SRAS-CoV, se evalúa su capacidad para inducir una respuesta inmunitaria protectora contra el SRAS-CoV en el ratón CD46^{+/-} IFN- α β R^{-/-}, que es susceptible a la infección por el MV. La respuesta en anticuerpos de los ratones inmunizados se evalúa mediante ensayo ELISA contra los antígenos nativos del SRAS-CoV y para su capacidad para neutralizar la infectividad del SRAS-CoV *in vitro*, utilizando las metodologías descritas anteriormente. El poder protector de la respuesta se evaluará midiendo la reducción de la carga vírica pulmonar 2 días después de una infección de prueba no letal por el SRAS-CoV.

Se construyen virus de rubeola recombinantes de segunda generación sustituyendo las secuencias salvajes de los genes S y Ssol por genes sintéticos optimizados para la expresión en células de mamíferos, descritos en el ejemplo 15 anteriormente. Estos virus de rubeola recombinantes son susceptibles de expresar cantidades más importantes de antígenos S y Ssol y por lo tanto presentar una inmunogenecidad incrementada.

Alternativamente, los genes salvajes o sintéticos que codifican para la proteína S o el polipéptido Ssol pueden ser insertados en el vector de la rubeola MVSchw-ATU3 en forma de una unidad de transcripción suplementaria localizada entre los genes H y L, y después los virus recombinantes producidos y caracterizados de manera similar. Esta inserción es susceptible de generar unos virus recombinantes que poseen unas características (multiplicación del virus, nivel de expresión del transgen) diferentes, y posiblemente una inmunogenecidad mejorada con respecto a los obtenidos después de la inserción de los transgenes entre los genes P y N.

El virus de la rubeola recombinante MVSchw2-SRAS-Ssol se puede utilizar para la producción en cantidad y la purificación del antígeno Ssol en vista a aplicaciones de diagnósticos y vacvíneos.

Ejemplo 18: Otras aplicaciones relacionadas con la proteína S

a) Los vectores lentivirales que permiten la expresión de la S o de Ssol (incluso de fragmentos de la S) pueden constituir una vacuna recombinante contra el SRAS-CoV, para ser utilizado en profilaxis humana y veterinaria. A fin de demostrar la factibilidad de tal vacuna, se estudia en el ratón la inmunogenecidad de los vectores lentivirales recombinantes TRIP-SD/SA-S-WPRE y TRIP-SD/SA-Ssol-WPRE.

b) Unos anticuerpos monoclonales son producidos con la ayuda del polipéptido recombinante Ssol. Según los resultados representados en el ejemplo 14 anterior, estos anticuerpos o por lo menos la mayoría de ellos reconocerán la forma nativa de la S del SRAS-CoV y serán susceptibles de aplicaciones de diagnósticos y/o profilácticas.

c) Un ensayo de serología del SRAS se realiza con el polipéptido Ssol utilizado como antígeno y la metodología del doble epítipo.

LISTADO DE SECUENCIAS

<110> INSTITUT PASTEUR
CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE
UNIVERSITE PARIS 7
VAN DER WERF, Sylvie
ESCRIOU, Nicolas
CRESCENZO-CHAIGNE, Bernadette
MANUGUERRA, Jean-Claude
KUNST, Franck
CALLENDRET, Benoît
BETTON, Jean-Michel
LORIN, Valérie
GERBAUD, Sylvie
BURGUIERE, Ana Maria
AZEBI, Saliha
CHARNEAU, Pierre
TANGY, Frédéric

COMBREDET, Chantai
DELAGNEAU, Jean-François
MARTIN, Monique

<120> UTILIZACIÓN DE LAS PROTEÍNAS Y DE LOS PÉPTIDOS CODIFICADOS POR EL GENOMA DE UNA
NUEVA CEPA DE CORONAVIRUS ASOCIADO AL SRAS

<130> 226-111ext

<150> FR 0314152

<151> 02-12-2003

<150> FR 0314151

<151> 02-12-2003

<160> 158

<170> PatentIn version 3.1

<210> 1

<211> 29746

<212> ADN

<213> CORONAVIRUS

<400> 1

atattagggtt	tttacctacc	caggaaaagc	caaccaacct	cgatctcttg	tagatctggt	60
ctctaaacga	actttaaaat	ctgtgtagct	gtcgctcggc	tgcatgccta	gtgcacctac	120
gcagtataaa	caataataaa	ttttactgtc	gttgacaaga	aacgagtaac	tcgtccctct	180
tctgcagact	gcttacgggt	tcgtccgtgt	tgcagtcgat	catcagcata	cctagggttc	240
gtccgggtgt	gaccgaaagg	taagatggag	agccttggtc	ttggtgtcaa	cgagaaaaca	300
cacgtccaac	tcagtttgcc	tgtccttcag	gttagagacg	tgctagtgcg	tggcttcggg	360
gactctgtgg	aagaggccct	atcggaaggc	cgtgaacacc	tcaaaaatgg	cacttggtgt	420
ctagtagagc	tggaaaaagg	cgtactgccc	cagcttgaac	agccctatgt	gttcattaaa	480
cgttctgatg	ccttaagcac	caatcacggc	cacaaggctg	ttgagctggt	tgcagaaatg	540
gacggcattc	agtacggctg	tagcgggtata	acactgggag	tactcgtgcc	acatgtgggc	600
gaaaccccaa	ttgcataccg	caatgttctt	cttcgtaaga	acggtaataa	gggagccggt	660
ggtcatagct	atggcatcga	tctaaagtct	tatgacttag	gtgacgagct	tggcactgat	720
cccattgaag	attatgaaca	aaactggaac	actaagcatg	gcagtgggtg	actccgtgaa	780
ctcactcgtg	agctcaatgg	aggtgcagtc	actcgtctatg	tcgacaacaa	tttctgtggc	840
ccagatgggt	accctcttga	ttgcatcaaa	gattttctcg	cacgcgcggg	caagtcaatg	900
tgcactcttt	ccgaacaact	tgattacatc	gagtcgaaga	gaggtgtcta	ctgctgccgt	960
gaccatgagc	atgaaattgc	ctggttcact	gagcgtctctg	ataagagcta	cgagcaccag	1020
acacccttcg	aaattaagag	tgccaagaaa	tttgacactt	tcaaagggga	atgcccaaag	1080
tttgtgtttc	ctcttaactc	aaaagtcaaa	gtcattcaac	cacgtgttga	aaagaaaaag	1140
actgaggggt	tcatggggcg	tatacgctct	gtgtaccctg	ttgcatctcc	acaggagtgt	1200
aacaatatgc	acttgctctac	cttgatgaaa	tgtaatcatt	gcgatgaagt	ttcatggcag	1260
acgtgcgact	ttctgaaagc	cacttgtgaa	cattgtggca	ctgaaaattt	agttattgaa	1320
ggacctacta	catgtgggta	cctacctact	aatgctgtag	tgaaaatgcc	atgtcctgcc	1380
tgtcaagacc	cagagattgg	acctgagcat	agtgttgtag	attatcacia	ccactcaaac	1440
attgaaactc	gactccgcaa	gggaggtagg	actagatggt	ttggaggctg	tgtgtttgcc	1500
tatgttggct	gctataataa	gcgtgcctac	tgggttcctc	gtgctagtgc	tgatattggc	1560

tcaggccata	ctggcattac	tggtgacaat	gtggagacct	tgaatgagga	tctccttgag	1620
atactgagtc	gtgaacgtgt	taacattaac	attgttggcg	attttcatit	gaatgaagag	1680
gttgccatca	ttttggcatc	tttctctgct	tctacaagtg	cctttattga	cactataaag	1740
agtcttgatt	acaagtcttt	caaaaccatt	gttgagtcct	gcggtaaacta	taaagttacc	1800
aagggaaagc	ccgtaaaagg	tgcttggaac	attggacaac	agagatcagt	tttaacacca	1860
ctgtgtggtt	ttccctcaca	ggctgctggt	gttatcagat	caatttttgc	gcgcacactt	1920
gatgcagcaa	accactcaat	tcctgatttg	caaagagcag	ctgtcaccat	acttgatggt	1980
atttctgaac	agtcattacg	tcttgctgac	gccatggttt	atacttcaga	cctgctcacc	2040
aacagtgtca	ttattatggc	atatgtaact	ggtggctcttg	tacaacagac	ttctcagtgg	2100
ttgtctaate	ttttgggcac	tactgttgaa	aaactcaggc	ctatctttga	atggattgag	2160
gcgaaactta	gtgcaggagt	tgaatttctc	aaggatgctt	gggagattct	caaatttctc	2220
attacaggtg	tttttgacat	cgtcaagggg	caaatacagg	ttgcttcaga	taacatcaag	2280
gattgtgtaa	aatgcttcat	tgatgttggt	aacaaggcac	tcgaaatgtg	cattgatcaa	2340
gtcactatcg	ctggcgcaaa	gttgcgatca	ctcaacttag	gtgaagtctt	catcgctcaa	2400
agcaagggac	tttaccgtca	gtgtatacgt	ggcaaggagc	agctgcaact	actcatgcct	2460
cttaaggcac	caaaagaagt	aacctttctt	gaagggtgatt	cacatgacac	agtacttacc	2520
tctgaggagg	ttgttctcaa	gaacggtgaa	ctcgaagcac	tcgagacgcc	cgttgatagc	2580
ttcaciaaatg	gagctatcgt	tggcacacca	gtctgtgtaa	atggcctcat	gctcttagag	2640
attaaggaca	aagaacaata	ctgcgcattg	tctcctgggt	tactggctac	aaacaatgtc	2700
tttcgcttaa	aaggggggtg	accaattaaa	ggtgtaacct	ttggagaaga	tactgtttgg	2760
gaagttcaag	gttacaagaa	tgtgagaatc	acatttgagc	ttgatgaacg	tgttgacaaa	2820
gtgcttaatg	aaaagtgtc	tgtctacact	gttgaatccg	gtaccgaagt	tactgagttt	2880
gcatgtgttg	tagcagaggc	tggtgtgaag	actttacaac	cagtttctga	tctccttacc	2940
aacatgggta	ttgatcttga	tgagtggagt	gtagctacat	tctacttatt	tgatgatgct	3000
ggtgaagaaa	acttttcatc	acgtatgtat	tgttcctttt	accctccaga	tgaggaagaa	3060
gaggacgatg	cagagtgtga	ggaagaagaa	attgatgaaa	cctgtgaaca	tgagtacggg	3120
acagaggatg	attatcaagg	tctccctctg	gaatttggtg	cctcagctga	aacagttcga	3180
gttgaggaag	aagaagagga	agactggctg	gatgatacta	ctgagcaatc	agagattgag	3240
ccagaaccag	aacctacacc	tgaagaacca	gttaatcagt	ttactgggtta	tttaaaactt	3300
actgacaatg	ttgccattaa	atgtgttgac	atcgtaagg	aggcacaag	tgctaatcct	3360
atgggtgattg	taaatgctgc	taacatacac	ctgaaacatg	gtgggtgggtg	agcagggtgca	3420
ctcaacaagg	caaccaatgg	tgccatgcaa	aaggagagtg	atgattacat	taagctaaat	3480
ggccctctta	cagtaggagg	gtcttggttg	ctttctggac	ataatcttgc	taagaagtgt	3540
ctgcatgttg	ttggacctaa	cctaaatgca	ggtgaggaca	tccagcttct	taaggcagca	3600

tatgaaaatt	tcaattcaca	ggacatctta	cttgaccat	tgttgtcagc	aggcatat	3660
ggtgctaaac	cacttcagtc	tttacaagt	tgcgtgcaga	cggttcgtac	acaggtttat	3720
attgcagtc	atgacaaagc	tctttatgag	caggttgtca	tggattatct	tgataacctg	3780
aagcctagag	tggaagcacc	taaacaagag	gagccaccaa	acacagaaga	ttccaaaact	3840
gaggagaaat	ctgtcgtaca	gaagcctgtc	gatgtgaagc	caaaaattaa	ggcctgcatt	3900
gatgagggtta	ccacaacact	ggaagaaact	aagtttctta	ccaataagtt	actcttgttt	3960
gctgatatca	atggtaagct	ttaccatgat	tctcagaaca	tgcttagagg	tgaagatatg	4020
tctttccttg	agaaggatgc	accttacatg	gtagggtgatg	ttatcactag	tggtgatatc	4080
acttggtgtg	taataccctc	caaaaaggct	ggtggcacta	ctgagatgct	ctcaagagct	4140
ttgaagaaag	tgccagttga	tgagtatata	accacgtacc	ctggacaagg	atgtgctggt	4200
tatacacttg	aggaagctaa	gactgctctt	aagaaatgca	aatctgcatt	ttatgtacta	4260
ccttcagaag	cacctaattgc	taaggaagag	attctaggaa	ctgtatcctg	gaatttgaga	4320
gaaatgcttg	ctcatgctga	agagacaaga	aaattaatgc	ctatatgcat	ggatgttaga	4380
gccataatgg	caaccatcca	acgtaagtat	aaaggaatta	aaattcaaga	gggcatcggt	4440
gactatggtg	tccgattctt	cttttatact	agtaaagagc	ctgtagcttc	tattattacg	4500
aagctgaact	ctctaaatga	gccgcttgct	acaatgccaa	ttggttatgt	gacacatggt	4560
tttaatcttg	aagaggctgc	gcgctgtatg	cgttctctta	aagctcctgc	cgtagtgctca	4620
gtatcatcac	cagatgctgt	tactacatat	aatggatacc	tcacttcgtc	atcaaagaca	4680
tctgaggagc	actttgtaga	aacagtttct	ttggctggct	cttacagaga	ttggctctat	4740
tcaggacagc	gtacagagtt	aggtgttgaa	tttcttaagc	gtggtgacaa	aattgtgtac	4800
cacactctgg	agagccccgt	cgagtttcat	cttgacggtg	aggttctttc	acttgacaaa	4860
ctaaagagtc	tcttatccct	gcgggaggtt	aagactataa	aagtgttcac	aactgtggac	4920
aacactaatc	tccacacaca	gcttggtgat	atgtctatga	catatggaca	gcagtttgggt	4980
ccaacatact	tggtatggtgc	tgatgttaca	aaaattaaac	ctcatgtaaa	tcagaggggt	5040
aagactttct	ttgtactacc	tagtgatgac	acactacgta	gtgaagcttt	cgagtactac	5100
catactcttg	atgagagitt	tcttggtagg	tacatgtctg	ctttaaacca	cacaaagaaa	5160
tggaattttc	ctcaagttgg	tggtttaact	tcaattaaat	gggctgataa	caattgttat	5220
ttgtctagt	ttttattagc	acttcaacag	cttgaagtca	aattcaatgc	accagcactt	5280
caagaggctt	attatagagc	ccgtgctgggt	gatgctgcta	acttttgtgc	actcactctc	5340
gcttacagta	ataaaactgt	tggtgagctt	ggtgatgtca	gagaaactat	gacccatctt	5400
ctacagcatg	ctaatttgga	atctgcaaag	cgagttctta	atgtggtgtg	taaacattgt	5460
ggtcagaaaa	ctactacctt	aacgggtgta	gaagctgtga	tgtatatggg	tactctatct	5520
tatgataatc	ttaagacagg	tgtttccatt	ccatgtgtgt	gtggctcgtga	tgctacacaa	5580
tatctagtag	aacaagagtc	ttcttttgtt	atgatgtctg	caccacctgc	tgagtataaa	5640

ttacagcaag	gtacattctt	atgtgcgaat	gagtacactg	gtaactatca	gtgtgggtcat	5700
tacactcata	taactgctaa	ggagaccctc	tatcgtattg	acggagctca	ccttacaaag	5760
atgtcagagt	acaaaggacc	agtgactgat	gttttctaca	aggaaacatc	ttacactaca	5820
accatcaagc	ctgtgtcgta	taaactcgat	ggagttactt	acacagagat	tgaacccaaa	5880
ttggatgggt	attataaaaa	ggataatgct	tactatacag	agcagcctat	agaccttgta	5940
ccaactcaac	cattaccaaa	tgcgagtttt	gataatttca	aactcacatg	ttctaacaca	6000
aaatttgctg	atgattttaa	tcaaatgaca	ggcttcacaa	agccagcttc	acgagagcta	6060
tctgtcacat	tcttcccgag	cttgaatggc	gatgtagtgg	ctattgacta	tagacactat	6120
tcagcgagtt	tcaagaaagg	tgctaaatta	ctgcataagc	caattgtttg	gcacattaac	6180
caggctacaa	ccaagacaac	gttcaaacca	aacacttggg	gtttacgttg	tctttggagt	6240
acaaagccag	tagatacttc	aaattcattt	gaagttctgg	cagtagaaga	cacacaagga	6300
atggacaatc	ttgcttgtga	aagtcaacaa	cccacctctg	aagaagtagt	ggaaaatcct	6360
accatacaga	aggaagtcac	agagtgtgac	gtgaaaacta	ccgaagtgtg	aggcaatgtc	6420
atacttaaac	catcagatga	aggtgttaaa	gtaacacaag	agttagggtca	tgaggatctt	6480
atggctgctt	atgtggaaaa	cacaagcatt	accattaaga	aacctaatga	gctttcacta	6540
gccttagggt	taaaaacaat	tgccactcat	ggatttgctg	caattaatag	tgttccttgg	6600
agtaaaattt	tggcttatgt	caaaccattc	ttaggacaag	cagcaattac	aacatcaaatt	6660
tgcgctaaga	gattagcaca	acgtgtgttt	aacaattata	tgcccttatgt	gtttacatta	6720
ttgttccaat	tgtgtacttt	tactaaaagt	accaattcta	gaattagagc	ttcactacct	6780
acaactattg	ctaaaaatag	tgtaagagt	gttgctaaat	tatgtttgga	tgccggcatt	6840
aattatgtga	agtcacccaa	attttctaaa	ttgttcacaa	tcgctatgtg	gctattgttg	6900
ttaagtattt	gcttaggttc	tctaattctgt	gtaactgctg	cttttggtgt	actcttatct	6960
aattttgggtg	ctccttctta	ttgtaatggc	gttagagaat	tgtatcttaa	ttcgtctaac	7020
gttactacta	tggatttctg	tgaaggttct	tttccttgca	gcatttgttt	aagtggatta	7080
gactcccttg	attcttatcc	agctcttgaa	accattcagg	tgacgatttc	atcgtacaag	7140
ctagacttga	caatttttagg	tctggccgct	gagtgggttt	tggcatatat	gttgttcaca	7200
aaattctttt	atttatttagg	tctttcagct	ataatgcagg	tgttcttttg	ctattttgct	7260
agtcatttca	tcagcaattc	ttggctcatg	tggtttatca	ttagtattgt	acaaatggca	7320
cccgtttctg	caatgggttag	gatgtacatc	ttctttgctt	ctttctacta	catatggaag	7380
agctatgttc	atatcatgga	tggttgcacc	tcttcgactt	gcatgatgtg	ctataagcgc	7440
aatcgtgcca	cacgcgttga	gtgtacaact	attgttaatg	gcatgaagag	atctttctat	7500
gtctatgcaa	atggaggccg	tggcttctgc	aagactcaca	attggaattg	tctcaattgt	7560
gacacatttt	gcactggtag	tacattcatt	agtgatgaag	ttgctcgtga	tttgtcactc	7620
cagtttaaaa	gaccaatcaa	ccctactgac	cagtcatcgt	atattgttga	tagtgttgct	7680

gtgaaaaatg	gcgcgcttca	cctctactrr	gacaaggctg	gtcaaaagac	ctatgagaga	7740
catccgctct	cccatTTTTgt	caatttagac	aatttgagag	ctaacaacac	taaagggttca	7800
ctgcctatta	atgtcatagt	ttttgatggc	aagtccaaat	gcgacgagtc	tgcttctaag	7860
tctgcttctg	tgtactacag	tcagctgatg	tgccaacctt	ttctgttgct	tgaccaagct	7920
cttgatatcag	acgttggaga	tagtactgaa	gtttccgtta	agatgtttga	tgcttatgtc	7980
gacacctttt	cagcaacttt	tagtgttcct	atggaaaaac	ttaaggcact	tgttgctaca	8040
gctcacagcg	agttagcaaa	gggtgtagct	ttagatgggtg	tcctttctac	attcgtgtca	8100
gctgcccgcg	aagggtgttg	tgataccgat	gttgacacaa	aggatgttat	tgaatgtctc	8160
aaactttcac	atcactctga	cttagaagtg	acaggtgaca	gttgtaacaa	tttcatgctc	8220
acctataata	aggttgaaaa	catgacgccc	agagatcttg	gcgcatgtat	tgactgtaat	8280
gcaaggcata	tcaatgccca	agtagcaaaa	agtcacaatg	tttcaactcat	ctggaatgta	8340
aaagactaca	tgtctttatc	tgaacagctg	cgtaaacaaa	ttcgtagtgc	tgccaagaag	8400
aacaacatac	cttttagact	aacttgtgct	acaactagac	aggttgtcaa	tgtcataact	8460
actaaaatct	cactcaaggg	tggttaagatt	gttagtactt	gttttaaact	tatgcttaag	8520
gccacattat	tgtgcgttct	tgctgcattg	gtttgttata	tcgttatgcc	agtacataca	8580
ttgtcaatcc	atgatgggtt	cacaaatgaa	atcattgggt	acaaagccat	tcaggatggg	8640
gtcactcgty	acatcatttc	tactgatgat	tgTTTTgcaa	ataaacatgc	tggttttgac	8700
gcatggttta	gccagcgtgg	tggttcatac	aaaaatgaca	aaagctgccc	tgtagtagct	8760
gctatcatta	caagagagat	tggtttcata	gtgcctggct	taccgggtac	tgtgctgaga	8820
gcaatcaatg	gtgacttctt	gcattttcta	cctcgtgttt	ttagtgctgt	tggcaacatt	8880
tgctacacac	cttccaaact	cattgagtat	agtgattttg	ctacctctgc	ttgcgttctt	8940
gctgctgagt	gtacaatttt	taaggatgct	atgggcaaac	ctgtgccata	ttgttatgac	9000
actaatttgc	tagagggttc	tatttcttat	agtgaagctt	gtccagacac	tcgttatgtg	9060
cttatggatg	gttccatcat	acagtttcct	aacacttacc	tggaggggtc	tgttagagta	9120
gtaacaactt	ttgatgctga	gtactgtaga	catggtacat	gcgaaaggct	agaagtaggt	9180
atttgcctat	ctaccagtgg	tagatgggtt	cttaataatg	agcattacag	agctctatca	9240
ggagttttct	gtggtgttga	tgcatgaat	ctcatagcta	acatctttac	tcctcttgty	9300
caacctgtgg	gtgctttaga	tgtgtctgct	tcagtagtgg	ctggtgggtat	tattgccata	9360
ttggtgactt	gtgctgccta	ctactttatg	aaattcagac	gtgtttttgg	tgagtacaac	9420
catgttgtyg	ctgctaattg	acttttgttt	ttgatgtctt	tcactatact	ctgtctggta	9480
ccagcttaca	gctttctgcc	gggagtctac	tcagtctttt	acttgacttt	gacattctat	9540
ttcaccaatg	atgtttcatt	cttggtctac	cttcaatggg	ttgccatggt	ttctcctatt	9600
gtgccttttt	ggataacagc	aatctatgta	ttctgtattt	ctctgaagca	ctgccattgg	9660
ttctttaaca	actatcttag	gaaaagagtc	atgtttaatg	gagttacatt	tagtaccttc	9720

gaggaggctg ctttgtgtac ctttttgctc aacaaggaaa tgtacctaaa attgcgtagc 9780
gagacactgt tgccacttac acagtataac aggtatcttg ctctatataa caagtacaag 9840
tatttcagtg gagccttaga tactaccagc tatcgtgaag cagcttgctg ccacttagca 9900
aaggctctaa atgacttttag caactcaggt gctgatgttc tctaccaacc accacagaca 9960
tcaatcactt ctgctgttct gcagagtggg tttaggaaaa tggcattccc gtcaggcaaa 10020
gttgaagggt gcatgggtaca agtaacctgt ggaactacaa ctcttaattg attgtgggtg 10080
gatgacacag tatactgtcc aagacatgtc atttgcacag cagaagacat gcttaatcct 10140
aactatgaag atctgctcat tcgcaaatcc aaccatagct ttcttggtca ggctggcaat 10200
gttcaacttc gtgttattgg ccattctatg caaaattgtc tgcttaggct taaagttgat 10260
acttctaacc ctaagacacc caagtataaa tttgtccgta tccaacctgg tcaaacattt 10320
tcagttctag catgctacaa tggttcacca tctgggtgtt atcagtgtgc catgagacct 10380
aatcatacca ttaaagggtc tttccttaat ggatcatgtg gtagtggttg ttttaacatt 10440
gattatgatt gcgtgtcttt ctgctatatg catcatatgg agcttccaac aggagtacac 10500
gctgggtactg acttagaagg taaattctat ggtccatttg ttgacagaca aactgcacag 10560
gctgcaggta cagacacaac cataacatta aatgttttgg catggctgta tgctgctgtt 10620
atcaatgggtg atagggtggt tcttaataga ttcaccacta ctttgaatga ctttaacctt 10680
gtggcaatga agtacaacta tgaacctttg acacaagatc atgttgacat attgggacct 10740
ctttctgctc aaacaggaat tgccgtctta gatatgtgtg ctgctttgaa agagctgctg 10800
cagaatggta tgaatggtcg tactatcctt ggtagcacta ttttagaaga tgagtttaca 10860
ccatttgatg ttgtagaca atgctctggg gttaccttcc aaggtaagtt caagaaaatt 10920
gttaagggca ctcatcattg gatgctttta actttcttga catcactatt gattcttggt 10980
caaagtacac agtggtcact gtttttcttt gtttacgaga atgctttctt gccatttact 11040
cttgggtatta tggcaattgc tgcagtgtgct atgctgcttg ttaagcataa gcacgcattc 11100
ttgtgcttgt ttctgttacc ttctcttgca acagttgctt actttaatat ggtctacatg 11160
cctgctagct ggggtgatgcg tatcatgaca tggcttgaat tggctgacac tagcttgtct 11220
ggttataggc ttaaggattg tgttatgtat gcttcagctt tagttttgct tattctcatg 11280
acagctcgca ctgtttatga tgatgctgct agacgtgttt ggacactgat gaatgtcatt 11340
acacttgttt acaaagtcta ctatggtaat gcttttagatc aagctatttc catgtgggcc 11400
ttagttattt ctgtaacctc taactattct ggtgtcgtta cgactatcat gtttttagct 11460
agagctatag tgtttgtgtg tgttgagtat taccattgtt tatttattac tggcaacacc 11520
ttacagtgtg tcatgcttgt ttattgtttc ttaggctatt gttgctgctg ctactttggc 11580
cttttctgtt tactcaaccg ttacttcagg ctactcttg gtgtttatga ctacttggtc 11640
tctacacaag aatttaggta tatgaactcc caggggcttt tgcctcctaa gagtagtatt 11700
gatgctttca agcttaacat taagttgttg ggtattggag gtaaaccatg tatcaagggt 11760

gctactgtac agtctaaaat gtctgacgta aagtgcacat ctgtggtact gctctcggtt 11820
cttcaacaac ttagagtaga gtcattcttct aaattgtggg cacaatgtgt acaactccac 11880
aatgatattc ttcttgcaaa agacacaact gaagctttcg agaagatggg ttctcttttg 11940
tctgttttgc tatccatgca ggggtgctgta gacattaata ggttgtgcga ggaaatgctc 12000
gataaccgtg ctactcttca ggctattgct tcagaattta gttctttacc atcatatgcc 12060
gcttatgcca ctgcccagga ggcctatgag caggctgtag ctaatgggta ttctgaagtc 12120
gttctcaaaa agttaaagaa atctttgaat gtggctaaat ctgagtttga ccgtgatgct 12180
gccatgcaac gcaagttgga aaagatggca gatcaggcta tgacccaaat gtacaaacag 12240
gcaagatctg aggacaagag ggcaaaagta actagtgcga tgcaaacaat gctcttcact 12300
atgcttagga agcttgataa tgatgcactt aacaacatta tcaacaatgc gcgtgatggg 12360
tgtgttccac tcaacatcat accattgact acagcagcca aactcatggg tgttgtccct 12420
gattatggta cctacaagaa cacttgtgat ggtaacacct ttacatatgc atctgcactc 12480
tgggaaatcc agcaagttgt tgatgcggat agcaagattg ttcaacttag tgaaattaac 12540
atggacaatt caccaaattt ggcttggcct cttattgtta cagctctaag agccaactca 12600
gctgttaaac tacagaataa tgaactgagt ccagtagcac tacgacagat gtcctgtgcg 12660
gctggtacca cacaacagc ttgtactgat gacaatgcac ttgcctacta taacaattcg 12720
aaggagggta ggtttgtgct ggcattacta tcagaccacc aagatctcaa atgggctaga 12780
ttccctaaga gtgatggtag aggtacaatt tacacagaac tggaaccacc ttgtaggttt 12840
gttacagaca caccaaaagg gcctaaagtg aaatacttgt acttcatcaa aggcctaaac 12900
aacctaaata gaggtatggg gctgggcagt ttagctgcta cagtacgtct tcaggctgga 12960
aatgctacag aagtacctgc caattcaact gtgctttcct tctgtgcttt tgcagtagac 13020
cctgctaaag catataagga ttacctagca agtgaggagc aaccaatcac caactgtgtg 13080
aagatgttgt gtacacacac tggtagagga caggcaatta ctgtaacacc agaagctaac 13140
atggaccaag agtccttttg tgggtgcttca tgttgtctgt attgtagatg ccacattgac 13200
catccaaatc ctaaaggatt ctgtgacttg aaaggtaagt acgtccaaat acctaccact 13260
tgtgctaattg acccagtggg ttttacactt agaaacacag tctgtaccgt ctgcggaatg 13320
tggaagggtt atggctgtag ttgtgaccaa ctccgcgaac ccttgatgca gtctgcggat 13380
gcatcaacgt ttttaaacgg gtttgcgggtg taagtgcagc ccgtcttaca ccgtgcggca 13440
caggcactag tactgatgtc gtctacaggg cttttgatat ttacaacgaa aaagttgctg 13500
gttttgcaaa gttcctaaaa actaattgct gtcgcttcca ggagaaggat gaggaaggca 13560
atttattaga ctcttacttt gtagttaaga ggcatactat gtctaactac caacatgaag 13620
agactattta taacttggtt aaagattgtc cagcgggtgc tgtccatgac tttttcaagt 13680
ttagagtaga tgggtgacatg gtaccacata tatcacgtca gcgtctaact aaatacacia 13740
tggctgattt agtctatgct ctacgtcatt ttgatgaggg taattgtgat acattaaaag 13800

aaatactcgt	cacatacaat	tgctgtgatg	atgattat	caataagaag	gattgggatg	13860
acttcgtaga	gaatcctgac	atcttacg	tatatgctaa	cttaggtgag	cgtgtacgcc	13920
aatcattatt	aaagactgta	caattctg	atgctatg	tgatgcaggc	attgtagg	13980
tactgacatt	agataatcag	gatcttaatg	ggaactggta	cgatttcggt	gatttcgtac	14040
aagtagcacc	aggctgcgga	gttcctattg	tggattcata	ttactcattg	ctgatgcccc	14100
tcctcacttt	gactagggca	ttggctgctg	agtcccatat	ggatgctgat	ctcgcaaaac	14160
cacttattaa	gtgggatttg	ctgaaatatg	attttacgga	agagagactt	tgtctcttcg	14220
accgttattt	taaatattgg	gaccagacat	accatcccaa	ttgtattaac	tgtttggtatg	14280
ataggtgtat	ccttcattgt	gcaaacttta	atgtgttatt	ttctactgtg	ttccaccta	14340
caagtttttg	accactagta	agaaaaatat	ttgtagatgg	tgttcctttt	gttgtttcaa	14400
ctggatacca	ttttcgtgag	ttaggagtcg	tacataatca	ggatgtaa	ttacatagct	14460
cgcgctcag	tttcaaggaa	cttttagtgt	atgctgctga	tccagctatg	catgcagctt	14520
ctggcaattt	attgctagat	aaacgcacta	catgcttttc	agtagctgca	ctaacaaca	14580
atgttgcttt	tcaaactgtc	aaacccggta	attttaataa	agacttttat	gactttgctg	14640
tgtctaaagg	tttctttaag	gaaggaagtt	ctgttgaact	aaaacacttc	ttctttgctc	14700
aggatggcaa	cgctgctatc	agtgattatg	actattatcg	ttataatctg	ccaacaatgt	14760
gtgatatcag	acaactccta	ttcgtagttg	aagttgttga	taaatacttt	gattgttacg	14820
atggtggctg	tattaatgcc	aaccaagtaa	tcgttaacaa	tctggataaa	tcagctgggt	14880
tcccatTTAA	taaatgggg	aaggctagac	tttattatga	ctcaatgagt	tatgaggatc	14940
aagatgcact	tttcgcgtat	actaagcgta	atgtcatccc	tactataact	caaatagaatc	15000
ttagtatgc	cattagtgc	aagaatagag	ctcgcacctg	agctggtgtc	tctatctgta	15060
gtactatgac	aaatagacag	tttcatcaga	aattattgaa	gtcaatagcc	gccactagag	15120
gagctactgt	ggtaattgga	acaagcaagt	tttacgggtg	ctggcataat	atgttaaaaa	15180
ctgtttacag	tgatgtagaa	actccacacc	ttatgggttg	ggattatcca	aaatgtgaca	15240
gagccatgcc	taacatgctt	aggataatgg	cctctcttgt	tcttgctcgc	aaacataaca	15300
cttgctgtaa	cttatcacac	cgtttctaca	ggtagctaa	cgagtgtg	caagtattaa	15360
gtgagatgg	catgtgtggc	ggctcactat	atgttaaacc	aggtggaaca	tcatccggtg	15420
atgctacaac	tgcttatgct	aatagtgtct	ttaacatttg	tcaagctgtt	acagccaatg	15480
taaatgcact	tctttcaact	gatggtaata	agatagctga	caagtatgtc	cgcaatctac	15540
aacacaggct	ctatgagtgt	ctctatagaa	atagggatgt	tgatcatgaa	ttcgtggatg	15600
agttttacgc	ttacctgcgt	aaacatttct	ccatgatgat	tctttctgat	gatgccgttg	15660
tgtgctataa	cagtaactat	gcggctcaag	gtttagtagc	tagcattaag	aactttaagg	15720
cagttcttta	ttatcaaaat	aatgtgttca	tgtctgaggc	aaaatgttgg	actgagactg	15780
accttactaa	aggacctcac	gaattttgct	cacagcatac	aatgctagtt	aaacaaggag	15840

atgattacgt	gtacctgcct	tacccagatc	catcaagaat	attagggcga	ggctgttttg	15900
tcgatgatat	tgtcaaaaca	gatggtacac	ttatgattga	aagggttcgtg	tacttggtcta	15960
ttgatgctta	cccacttaca	aaacatccta	atcaggagta	tgctgatgtc	tttcacttgt	16020
atttacaata	cattagaaag	ttacatgatg	agcttactgg	ccacatgttg	gacatgtatt	16080
ccgtaatgct	aactaatgat	aacacctcac	ggtactggga	acctgagttt	tatgaggcta	16140
tgtacacacc	acatacagtc	ttgcaggctg	taggtgcttg	tgtattgtgc	aattcacaga	16200
cttcacttcg	ttgcggtgcc	tgtattagga	gaccattcct	atgttgcaag	tgctgctatg	16260
accatgtcat	ttcaacatca	cacaaattag	tgttgctctgt	taatccctat	gtttgcaatg	16320
ccccaggttg	tgatgtcact	gatgtgacac	aactgtatct	aggaggtagt	agctattatt	16380
gcaagtcaca	taagcctccc	attagttttc	cattatgtgc	taatggtcag	gtttttgggt	16440
tatacaaaaa	cacatgtgta	ggcagtgaca	atgtcactga	cttcaatgcg	atagcaacat	16500
gtgattggac	taatgctggc	gattacatac	ttgccaacac	ttgtactgag	agactcaagc	16560
ttttcgcagc	agaaacgctc	aaagccactg	aggaaacatt	taagctgtca	tatgggtattg	16620
ccactgtacg	cgaagtactc	tctgacagag	aattgcatct	ttcatgggag	gttggaatac	16680
ctagaccacc	attgaacaga	aactatgtct	ttactgggta	ccgtgtaact	aaaaatagta	16740
aagtacagat	tggagagtag	acctttgaaa	aagggtgacta	tgggtgatgct	gttgtgtaca	16800
gagggtactac	gacatacaag	ttgaatgttg	gtgattactt	tgtgttgaca	tctcacactg	16860
taatgccact	tagtgcacct	actctagtgc	cacaagagca	ctatgtgaga	attactggct	16920
tgtacccaac	actcaacatc	tcagatgagt	tttctagcaa	tggttgcaat	tatcaaaagg	16980
tcggcatgca	aaagtactct	acactccaag	gaccacctgg	tactggtaag	agtcattttg	17040
ccatcggact	tgctctctat	tacccatctg	ctcgcatagt	gtatacggca	tgctctcatg	17100
cagctgttga	tgccctatgt	gaaaaggcat	taaaatattt	gcccatagat	aaatgtagta	17160
gaatcatacc	tgcgctgctg	cgcgtagagt	gttttgataa	attcaaagtg	aattcaacac	17220
tagaacagta	tgttttctgc	actgtaaatg	cattgccaga	aacaactgct	gacattgtag	17280
tctttgatga	aatctctatg	gctactaatt	atgacttgag	tggtgtcaat	gctagacttc	17340
gtgcaaaaaca	ctacgtctat	attggcgatc	ctgctcaatt	accagcccc	cgcacattgc	17400
tgactaaagg	cacactagaa	ccagaatatt	ttaattcagt	gtgcagactt	atgaaaacaa	17460
taggtccaga	catgttcctt	ggaacttgct	gccgttgctc	tgctgaaatt	gttgacactg	17520
tgagtgcitt	agtttatgac	aataagctaa	aagcacacaa	ggataagtca	gctcaatgct	17580
tcaaaatggt	ctacaaagg	gttattacac	atgatgtttc	atctgcaatc	aacagacctc	17640
aaataggcgt	tgtaagagaa	tttcttacac	gcaatcctgc	ttggagaaaa	gctgttttta	17700
tctcacctta	taattcacag	aacgctgtag	cttcaaaaat	cttaggattg	cctacgcaga	17760
ctgttgattc	atcacagggt	tctgaatatg	actatgtcat	attcacacaa	actactgaaa	17820
cagcacactc	ttgtaatgtc	aaccgcttca	atgtggctat	cacaagggca	aaaattggca	17880

ttttgtgcat	aatgtctgat	agagatcttt	atgacaaact	gcaatttaca	agtctagaaa	17940
taccacgtcg	caatgtggct	acattacaag	cagaaaatgt	aactggactt	tttaaggact	18000
gtagtaagat	cattactggg	cttcataccta	cacaggcacc	tacacacctc	agcgttgata	18060
taaagttcaa	gactgaagga	ttatgtgttg	acataccagg	cataccaaag	gacatgacct	18120
accgtagact	catctctatg	atgggtttca	aaatgaatta	ccaagtcaat	ggttacccta	18180
atatgtttat	cacccgcgaa	gaagctattc	gtcacgttcg	tgcgtggatt	ggctttgatg	18240
tagagggctg	tcatgcaact	agagatgctg	tgggtactaa	cctacctctc	cagctaggat	18300
tttctaagc	tgtaactta	gtagctgtac	cgactgggta	tgttgacact	gaaaataaca	18360
cagaattcac	cagagttaat	gcaaaacctc	caccagggtga	ccagttaaaa	catcttatac	18420
cactcatgta	taaaggcttg	ccctggaatg	tagtgcgat	taagatagta	caaagtctca	18480
gtgatacact	gaaaggattg	tcagacagag	tcgtgttcgt	cctttgggcg	catggctttg	18540
agcttacatc	aatgaagtac	tttgtcaaga	ttggacctga	aagaacgtgt	tgtctgtgtg	18600
acaaacgtgc	aacttgcttt	tctacttcat	cagatactta	tgcctgctgg	aatcattctg	18660
tgggttttga	ctatgtctat	aaccatttta	tgattgatgt	tcagcagtgg	ggctttacgg	18720
gtaaccttca	gagtaaccat	gaccaacatt	gccagggtaca	tggaaatgca	catgtggcta	18780
gttgtgatgc	tatcatgact	agatgttttag	cagtccatga	gtgctttgtt	aagcgcgttg	18840
attggtctgt	tgaataccct	attataggag	atgaactgag	ggtaattct	gcttgcagaa	18900
aagtacaaca	catggttgtg	aagtctgcat	tgcttgctga	taagtttcca	gttcttcatg	18960
acattggaaa	tccaaaggct	atcaagtgtg	tgcctcaggc	tgaagtagaa	tggaaagtct	19020
acgatgctca	gccatgtagt	gacaaagctt	acaaaataga	ggaactcttc	tattcttatg	19080
ctacacatca	cgataaattc	actgatgggtg	tttgtttgtt	ttggaattgt	aacgttgatc	19140
gttaccagc	caatgcaatt	gtgtgtaggt	ttgacacaag	agtcttgctca	aacttgaact	19200
taccaggctg	tgatgggtgg	agtttgatatg	tgaataagca	tgcattccac	actccagctt	19260
tcgataaaaag	tgcatttact	aattttaaagc	aattgccttt	cttttactat	tctgatagtc	19320
cttgtagatc	tcatggcaaa	caagtagtgt	cggatattga	ttatgttcca	ctcaaactctg	19380
ctacgtgtat	tacacgatgc	aatttaggtg	gtgctgtttg	cagacaccat	gcaaatgagt	19440
accgacagta	cttggaatgca	tataatatga	tgatttctgc	tggatttagc	ctatggattt	19500
acaaacaatt	tgatacttat	aacctgtgga	atacatttac	caggttacag	agtttagaaa	19560
atgtggctta	taatgttgtt	aataaaggac	actttgatgg	acacgccggc	gaagcacctg	19620
tttccatcat	taataatgct	gtttacacaa	aggtagatgg	tattgatgtg	gagatctttg	19680
aaaataagac	aacacttcct	gttaatgttg	catttgagct	ttgggctaag	cgtaacatta	19740
aaccagtgcc	agagattaag	atactcaata	atttggtgtg	tgatatcgct	gctaatactg	19800
taatctggga	ctacaaaaga	gaagccccag	cacatgtatc	tacaataggt	gtctgcacaa	19860
tgactgacat	tgccaagaaa	cctactgaga	gtgcttgctc	ttcacttact	gtcttggttg	19920

atggttagagt	ggaaggacag	gtagaccttt	ttagaaacgc	ccgtaatggt	gttttaataa	19980
cagaagggttc	agtcaaagggt	ctaacacctt	caaagggacc	agcacaagct	agcgtcaatg	20040
gagtcacatt	aattggagaa	tcagtaaaaa	cacagtttaa	ctactttaag	aaagtagacg	20100
gcattattca	acagttgcct	gaaacctact	ttactcagag	cagagactta	gaggatttta	20160
agcccagatc	acaaatggaa	actgactttc	tcgagctcgc	tatggatgaa	ttcatacagc	20220
gatataagct	cgagggctat	gccttcgaac	acatcgttta	tggagatttc	agtcattggac	20280
aacttggcgg	tcttcattta	atgataggct	tagccaagcg	ctcacaagat	tcaccactta	20340
aattagagga	ttttatccct	atggacagca	cagtgaaaaa	ttacttcata	acagatgcgc	20400
aaacagggttc	atcaaaatgt	gtgtgttctg	tgattgatct	tttacttgat	gactttgtcg	20460
agataataaa	gtcacaagat	ttgtcagtga	tttcaaaagt	ggccaagggt	acaattgact	20520
atgctgaaat	ttcattcatg	ctttggtgta	aggatggaca	tgttgaaacc	ttctacccaa	20580
aactacaagc	aagtcaagcg	tggcaaccag	gtgttgcatg	gcctaacttg	tacaagatgc	20640
aaagaatgct	tcttgaaaag	tgtgaccttc	agaattatgg	tgaaaatgct	gttataccaa	20700
aaggaataat	gatgaatgtc	gcaaagtata	ctcaactgtg	tcaatactta	aatacactta	20760
ctttagctgt	accctacaac	atgagagtta	ttcactttgg	tgctggctct	gataaaggag	20820
ttgcaccagg	tacagctgtg	ctcagacaat	ggttgccaac	tggcacacta	cttgtcgtat	20880
cagatcttaa	tgacttcgtc	tccgacgcag	attctacttt	aattggagac	tgtgcaacag	20940
tacatacggc	taataaatgg	gaccttatta	ttagcgatat	gtatgacctt	aggaccaaac	21000
atgtgacaaa	agagaatgac	tctaaagaag	ggtttttcac	ttatctgtgt	ggattttataa	21060
agcaaaaact	agccctgggt	ggttctatag	ctgtaaagat	aacagagcat	tcttggaatg	21120
ctgaccttta	caagcttatg	ggccatttct	catgggtggac	agcttttggt	acaaatgtaa	21180
atgcatcatc	atcggaagca	tttttaattg	gggctaacta	tcttggcaag	ccgaaggaac	21240
aaattgatgg	ctataccatg	catgctaact	acattttctg	gaggaacaca	aatcctatcc	21300
agttgtcttc	ctattcactc	tttgacatga	gcaaatttcc	tcttaaatta	agaggaactg	21360
ctgtaatgtc	tcttaaggag	aatcaaatca	atgatatgat	ttattctctt	ctggaaaaag	21420
gtaggcttat	cattagagaa	aacaacagag	ttgtggtttc	aagtgatatt	cttggttaaca	21480
actaaacgaa	catgtttatt	ttcttattat	ttcttactct	cactagtggg	agtgcacttg	21540
accggtgcac	cacttttgat	gatgttcaag	ctcctaatta	cactcaacat	acttcatcta	21600
tgaggggggt	ttactatcct	gatgaaattt	ttagatcaga	cactctttat	ttactcagg	21660
atttattttct	tccattttat	tctaattgta	caggggtttca	tactattaat	catacgtttg	21720
gcaaccctgt	catacctttt	aaggatggta	tttattttgc	tgccacagag	aaatcaaattg	21780
ttgtccgtgg	ttgggttttt	ggttctacca	tgaacaacaa	gtcacagtcg	gtgattatta	21840
ttaacaattc	tactaatgtt	gttatacgag	catgtaactt	tgaattgtgt	gacaaccctt	21900
tctttgctgt	ttctaaaccc	atgggtacac	agacacatac	tatgatattc	gataatgcat	21960

ttaattgcac tttcgagtac atatctgatg ccttttcgct tgatgtttca gaaaagtcag 22020
 gtaattttta acacttacga gagtttgtgt ttaaaaaataa agatggggtt ctctatgttt 22080
 ataagggcta tcaacctata gatgtagttc gtgatctacc ttctgggttt aacactttga 22140
 aacctatttt taagtgcct cttgggtatta acattacaaa ttttagagcc attcttacag 22200
 ccttttcacc tgctcaagac atttggggca cgtcagctgc agcctatttt gttggctatt 22260
 taaagccaac tacatttatg ctcaagtatg atgaaaatgg tacaatcaca gatgctgttg 22320
 attgttctca aaatccactt gctgaactca aatgctctgt taagagcttt gagattgaca 22380
 aaggaattta ccagacctct aatttcaggg ttgttccctc aggagatgtt gtgagattcc 22440
 ctaataattac aaacttgtgt ccttttgag aggtttttta tgctactaaa ttcccttctg 22500
 tctatgcatg ggagagaaaa aaaatttcta attgtgttgc tgattactct gtgctctaca 22560
 actcaacatt tttttcaacc ttttaagtgt atggcgtttc tgccactaag ttgaatgatc 22620
 tttgcttctc caatgtctat gcagattctt ttgtagtcaa gggagatgat gtaagacaaa 22680
 tagcgccagg acaaactggg gttattgctg attataatta taaattgcca gatgatttca 22740
 tgggttgtgt ccttgcttgg aatactagga acattgatgc tacttcaact ggtaattata 22800
 attataaata taggtatctt agacatggca agcttaggcc ctttgagaga gacatatcta 22860
 atgtgccttt ctcccctgat ggcaaacctt gcaccccacc tgctcttaat tgttattggc 22920
 cattaaatga ttatgggttt tacaccacta ctggcattgg ctaccaacct tacagagttg 22980
 tagtactttc ttttgaactt ttaaatgcac cggccacggg ttgtggacca aaattatcca 23040
 ctgaccttat taagaaccag tgtgtcaatt ttaattttta tggactcact ggtactgggtg 23100
 tgtaactcc ttcttcaaag agatttcaac catttcaaca atttggccgt gatgtttctg 23160
 atttcaactga ttccgttcga gatcctaaaa catctgaaat attagacatt tcaccttgc 23220
 cttttggggg tgtaagtgt attacacctg gaacaaatgc ttcatctgaa gttgctgttc 23280
 tatatcaaga tgtaactgc actgatgttt ctacagcaat tcatgcagat caactcacac 23340
 cagcttggcg catatattct actggaaaca atgtattcca gactcaagca ggctgtctta 23400
 taggagctga gcatgtcgac acttcttatg agtgcgacat tcctattgga gctggcattt 23460
 gtgctagtta ccatacagtt tctttattac gtagtactag ccaaaaatct attgtggctt 23520
 atactatgtc tttagggtgct gatagttcaa ttgcttactc taataacacc attgctatac 23580
 ctactaactt ttcaattagc attactacag aagtaatgcc tgtttctatg gctaaaacct 23640
 ccgtagattg taatatgtac atctgcggag attctactga atgtgctaatt ttgcttctcc 23700
 aatatggtag cttttgcaca caactaaatc gtgcactctc aggtattgct gctgaacagg 23760
 atcgcaacac acgtgaagtg ttcgctcaag tcaaacaaat gtacaaaacc ccaactttga 23820
 aatattttgg tggttttaat ttttcacaaa tattacctga ccctctaaag ccaactaaga 23880
 ggtcttttat tgaggacttg ctctttaata aggtgacact cgctgatgct ggcttcatga 23940
 agcaatatgg cgaatgccta ggtgatatta atgctagaga tctcatttgt gcgcagaagt 24000

tcaatggact	tacagtgttg	ccacctctgc	tcactgatga	tatgattgct	gcctacactg	24060
ctgctctagt	tagtggtact	gccactgctg	gatggacatt	tggtgctggc	gctgctcttc	24120
aaataccttt	tgctatgcaa	atggcatata	ggttcaatgg	cattggagtt	acccaaaatg	24180
ttctctatga	gaaccaaaaa	caaatcgcca	accaatttaa	caaggcgatt	agtcaaattc	24240
aagaatcact	tacaacaaca	tcaactgcat	tgggcaagct	gcaagacgtt	gttaaccaga	24300
atgctcaagc	attaaacaca	cttgttaaac	aacttagctc	taattttggt	gcaatttcaa	24360
gtgtgctaaa	tgatatcctt	tcgcgacttg	ataaagtcga	ggcggaggta	caaattgaca	24420
ggttaattac	aggcagactt	caaagccttc	aaacctatgt	aacacaacaa	ctaatacagg	24480
ctgctgaaat	cagggcttct	gctaatactg	ctgctactaa	aatgtctgag	tgtgttcttg	24540
gacaatcaaa	aagagttgac	ttttgtggaa	agggtacca	ccttatgtcc	ttcccacaag	24600
cagccccgca	tggtgttgtc	ttcctacatg	tcacgtatgt	gccatcccag	gagaggaact	24660
tcaccacagc	gccagcaatt	tgtcatgaag	gcaaagcata	cttccctcgt	gaagggtgtt	24720
ttgtgtttaa	tggcacttct	tggtttatta	cacagaggaa	cttcttttct	ccacaaataa	24780
ttactacaga	caatacattt	gtctcaggaa	attgtgatgt	cgttattggc	atcattaaca	24840
acacagttta	tgatcctctg	caacctgagc	ttgactcatt	caaagaagag	ctggacaagt	24900
acttcaaaaa	tcatacatca	ccagatgttg	atcttggcga	catttcaggc	attaacgctt	24960
ctgtcgtaaa	cattcaaaaa	gaaattgacc	gcctcaatga	ggtcgctaaa	aattttaaag	25020
aatcactcat	tgaccttcaa	gaattgggaa	aatatgagca	atatattaaa	tggccttggt	25080
atgtttggct	cggcttcatt	gctggactaa	ttgccatcgt	catgggtaca	atcttgcttt	25140
gttgcatgac	tagttgttgc	agttgcctca	agggtgcatg	ctcttggtgt	tcttgctgca	25200
agtttgatga	ggatgactct	gagccagttc	tcaagggtgt	caaattacat	tacacataaa	25260
cgaacttatg	gatttgttta	tgagattttt	tactcttgga	tcaattactg	cacagccagt	25320
aaaaattgac	aatgcttctc	ctgcaagtac	tgttcatgct	acagcaacga	taccgctaca	25380
agcctcactc	cctttcggat	ggcttggtat	tggcgttgca	tttcttgctg	tttttcagag	25440
cgctaccaaa	ataattgcgc	tcaataaaaag	atggcagcta	gccctttata	agggtctcca	25500
gttcatttgc	aatttactgc	tgctatttgt	taccatctat	tcacatcttt	tgcttgctgc	25560
tgagggtatg	gagggcgcaat	ttttgtacct	ctatgccttg	atatattttc	tacaatgcat	25620
caacgcatgt	agaattatta	tgagatgttg	gctttgttgg	aagtgcaa	ccaagaaccc	25680
attactttat	gatgccaa	actttgtttg	ctggcacaca	cataactatg	actactgtat	25740
accatataac	agtgtcacag	atacaattgt	cgttactgaa	ggtagcggca	tttcaacacc	25800
aaaactcaaa	gaagactacc	aaattgggtg	ttattctgag	gataggcact	caggtgttaa	25860
agactatgtc	gttgatcatg	gctatttcac	cgaagtgtac	taccagcttg	agtctacaca	25920
aattactaca	gacactggta	ttgaaaatgc	tacattcttc	atctttaaca	agcttggttaa	25980
agaccaccg	aatgtgcaaa	tacacacaat	cgacggctct	tcaggagttg	ctaataccagc	26040

aatggatcca	atztatgatg	agccgacgac	gactactagc	gtgcctttgt	aagcacaaga	26100
aagtgagtac	gaacttatgt	actcattcgt	ttcggaagaa	acagggtacgt	taatagttaa	26160
tagcgtactt	ctttttcttg	ctttcgtggt	attcttgcta	gtcacactag	ccatccttac	26220
tgcgcttcga	ttgtgtgcgt	actgctgcaa	tattgttaac	gtgagtttag	taaaaccaac	26280
ggtttacgtc	tactcgcgtg	ttaaaaatct	gaactcttct	gaaggagttc	ctgatcttct	26340
ggtctaaacg	aactaactat	tattattatt	ctgtttggaa	ctttaacatt	gcttatcatg	26400
gcagacaacg	gtactattac	cgttgaggag	cttaacaac	tcctggaaca	atggaaccta	26460
gtaatagggt	tcctattcct	agcctggatt	atgttactac	aatttgcta	ttctaatacg	26520
aacagggtttt	tgtacataat	aaagcttggt	ttcctctggc	tcttggtggc	agtaacactt	26580
gcttggtttg	tgcttgctgc	tgtctacaga	attaattggg	tgactggcgg	gattgagatt	26640
gcaatggctt	gtattgtagg	cttgatgtgg	cttagctact	tcgttgcttc	cttcaggctg	26700
tttgctcgta	cccgtcfaat	gtggtcattc	aaccagaaa	caaacattct	tctcaatgtg	26760
cctctccggg	ggacaattgt	gaccagaccg	ctcatggaaa	gtgaacttgt	cattggtgct	26820
gtgatcattc	gtggtcactt	gcgaatggcc	ggacactccc	tagggcgctg	tgacattaag	26880
gacctgccaa	aagagatcac	tgtggctaca	tcacgaacgc	tttcttatta	caaattagga	26940
gcgtcgagc	gtgtaggcac	tgattcaggt	tttgctgcat	acaaccgcta	ccgtattgga	27000
aactataaat	taaatacaga	ccacgccggt	agcaacgaca	atattgcttt	gctagtacag	27060
taagtgacaa	cagatgtttc	atcttggtga	cttcagggtt	acaatagcag	agatattgat	27120
tatcattatg	aggactttca	ggattgctat	ttggaatctt	gacgttataa	taagttcaat	27180
agtgagacaa	ttatttaagc	ctctaactaa	gaagaattat	tcggagttag	atgatgaaga	27240
acctatggag	ttagattatc	cataaaacga	acatgaaaat	tattctcttc	ctgacattga	27300
ttgtattttac	atcttgcgag	ctatatcact	atcaggagtg	tgttagaggt	acgactgtac	27360
tactaaaaga	accttgccca	tcaggaacat	acgagggcaa	ttcaccattt	cacctcttg	27420
ctgacaataa	atctgacta	acttgacta	gcacacactt	tgcttttgct	tgtgctgacg	27480
gtactcgaca	tacctatcag	ctgcgtgcaa	gatcagtttc	acaaaactt	ttcatcagac	27540
aagaggaggt	tcaacaagag	ctctactcgc	cactttttct	cattgttgct	gctctagtat	27600
ttttaatact	ttgcttcacc	attaagagaa	agacagaatg	aatgagctca	ctttaattga	27660
cttctatttg	tgctttttag	cctttctgct	attccttggt	ttaataatgc	ttattatatt	27720
ttggttttca	ctcgaaatcc	aggatctaga	agaaccttgt	accaaagtct	aaacgaacat	27780
gaaacttctc	attgttttga	cttgattttc	tctatgcagt	tgcatatgca	ctgtagtaca	27840
gcgtgtgca	tctaataaac	ctcatgtgct	tgaagatcct	tgtaaggtag	aacactaggg	27900
gtaatactta	tagcactgct	tggctttgtg	ctctaggaaa	ggttttacct	tttcatagat	27960
ggcacactat	ggttcaaaca	tgcacaccta	atgttactat	caactgtcaa	gatccagctg	28020
gtgggtgcgt	tatagctagg	tgttggtacc	ttcatgaagg	tcaccaaact	gctgcattta	28080

gagacgtact	tgttggttta	aataaacgaa	caaattaaaa	tgtctgataa	tggaccccaa	28140
tcaaaccaac	gtagtgcgcc	ccgcattaca	tttgggtggac	ccacagattc	aactgacaat	28200
aaccagaatg	gaggacgcaa	tggggcaagg	ccaaaacagc	gccgacccca	aggtttaccc	28260
aataatactg	cgtcttggtt	cacagctctc	actcagcatg	gcaaggagga	acttagattc	28320
cctcgaggcc	agggcgttcc	aatcaacacc	aatagtgggtc	cagatgacca	aattggctac	28380
taccgaagag	ctacccgacg	agttcgtggt	ggtgacggca	aaatgaaaga	gctcagcccc	28440
agatggtact	tctattacct	aggaactggc	ccagaagctt	cacttcacct	cggcgctaac	28500
aaagaaggca	tcgtatgggt	tgcaactgag	ggagccttga	atacacccaa	agaccacatt	28560
ggcaccgcga	atcctaataa	caatgctgcc	accgtgctac	aacttcctca	aggaacaaca	28620
ttgccaaaag	gcttctacgc	agaggggaagc	agaggcgga	gtcaagcctc	ttctcgctcc	28680
tcatacagta	gtcgcggtaa	ttcaagaaat	tcaactcctg	gcagcagtag	gggaaattct	28740
cctgctcgaa	tggctagcgg	aggtgggtgaa	actgccctcg	cgctattgct	gctagacaga	28800
ttgaaccagc	ttgagagcaa	agtttctggt	aaaggccaac	aacaacaagg	ccaaactgtc	28860
actaagaaat	ctgctgctga	ggcatctaaa	aagcctcgcc	aaaaacgtac	tgccacaaaa	28920
cagtacaacg	tcactcaagc	atgtgggaga	cgtggtccag	aacaaacca	aggaaatttc	28980
ggggaccaag	acctaattcag	acaaggaact	gattacaaac	attggccgca	aattgcacaa	29040
tttgctccaa	gtgcctctgc	attctttgga	atgtcacgca	ttggcatgga	agtcacacct	29100
tcgggaacat	ggctgactta	tcattggagcc	attaaattgg	atgacaaaga	tccacaattc	29160
aaagacaacg	tcatactgct	gaacaagcac	attgacgcat	acaaaacatt	cccaccaaca	29220
gagcctaaaa	aggacaaaaa	gaaaaagact	gatgaagctc	agcctttgcc	gcagagacaa	29280
aagaagcagc	ccactgtgac	tcttcttcct	gcggctgaca	tggatgattt	ctccagacaa	29340
cttcaaaatt	ccatgagtgg	agcttctgct	gattcaactc	aggcataaac	actcatgatg	29400
accacacaag	gcagatgggc	tatgtaaacg	ttttcgcaat	tccgtttacg	atacatagtc	29460
tactcttggtg	cagaatgaat	tctcgtaact	aaacagcaca	agtaggttta	gttaacttta	29520
atctcacata	gcaatcttta	atcaatgtgt	aacattaggg	aggacttgaa	agagccacca	29580
cattttcatc	gaggccacgc	ggagtacgat	cgaggggtaca	gtgaataatg	ctagggagag	29640
ctgcctatat	ggaagagccc	taatgtgtaa	aattaatttt	agtagtgcta	tccccatgtg	29700
attttaatag	cttcttagga	gaatgacaaa	aaaaaaaaaa	aaaaaa		29746

<210> 2
 <211>3945
 <212> ADN
 <213> CORONAVIRUS

<220>
 <221> CDS
 <222> (89)..(3853)
 <223>

ES 2 396 127 T3

<400> 2

ttctcttctg gaaaaaggta ggcttatcat tagagaaaac aacagagttg tggtttcaag	60
tgatattctt gttaacaact aaacgaac atg ttt att ttc tta tta ttt ctt	112
Met Phe Ile Phe Leu Leu Phe Leu	
1 5	
act ctc act agt ggt agt gac ctt gac cgg tgc acc act ttt gat gat	160
Thr Leu Thr Ser Gly Ser Asp Leu Asp Arg Cys Thr Thr Phe Asp Asp	
10 15 20	
gtt caa gct cct aat tac act caa cat act tca tct atg agg ggg gtt	208
Val Gln Ala Pro Asn Tyr Thr Gln His Thr Ser Ser Met Arg Gly Val	
25 30 35 40	
tac tat cct gat gaa att ttt aga tca gac act ctt tat tta act cag	256
Tyr Tyr Pro Asp Glu Ile Phe Arg Ser Asp Thr Leu Tyr Leu Thr Gln	
45 50 55	
gat tta ttt ctt cca ttt tat tct aat gtt aca ggg ttt cat act att	304
Asp Leu Phe Leu Pro Phe Tyr Ser Asn Val Thr Gly Phe His Thr Ile	
60 65 70	
aat cat acg ttt ggc aac cct gtc ata cct ttt aag gat ggt att tat	352
Asn His Thr Phe Gly Asn Pro Val Ile Pro Phe Lys Asp Gly Ile Tyr	
75 80 85	
ttt gct gcc aca gag aaa tca aat gtt gtc cgt ggt tgg gtt ttt ggt	400
Phe Ala Ala Thr Glu Lys Ser Asn Val Val Arg Gly Trp Val Phe Gly	
90 95 100	
tct acc atg aac aac aag tca cag tcg gtg att att att aac aat tct	448
Ser Thr Met Asn Asn Lys Ser Gln Ser Val Ile Ile Ile Asn Asn Ser	
105 110 115 120	
act aat gtt gtt ata cga gca tgt aac ttt gaa ttg tgt gac aac cct	496
Thr Asn Val Val Ile Arg Ala Cys Asn Phe Glu Leu Cys Asp Asn Pro	
125 130 135	
ttc ttt gct gtt tct aaa ccc atg ggt aca cag aca cat act atg ata	544
Phe Phe Ala Val Ser Lys Pro Met Gly Thr Gln Thr His Thr Met Ile	
140 145 150	
ttc gat aat gca ttt aat tgc act ttc gag tac ata tct gat gcc ttt	592
Phe Asp Asn Ala Phe Asn Cys Thr Phe Glu Tyr Ile Ser Asp Ala Phe	
155 160 165	
tcg ctt gat gtt tca gaa aag tca ggt aat ttt aaa cac tta cga gag	640
Ser Leu Asp Val Ser Glu Lys Ser Gly Asn Phe Lys His Leu Arg Glu	
170 175 180	
ttt gtg ttt aaa aat aaa gat ggg ttt ctc tat gtt tat aag ggc tat	688
Phe Val Phe Lys Asn Lys Asp Gly Phe Leu Tyr Val Tyr Lys Gly Tyr	
185 190 195 200	
caa cct ata gat gta gtt cgt gat cta cct tct ggt ttt aac act ttg	736
Gln Pro Ile Asp Val Val Arg Asp Leu Pro Ser Gly Phe Asn Thr Leu	
205 210 215	

aaa cct att ttt aag ttg cct ctt ggt att aac att aca aat ttt aga	784
Lys Pro Ile Phe 220 Lys Leu Pro Leu Gly 225 Ile Asn Ile Thr Asn 230 Phe Arg	
gcc att ctt aca gcc ttt tca cct gct caa gac att tgg ggc acg tca	832
Ala Ile Leu 235 Thr Ala Phe Ser Pro 240 Ala Gln Asp Ile Trp 245 Gly Thr Ser	
gct gca gcc tat ttt gtt ggc tat tta aag cca act aca ttt atg ctc	880
Ala Ala 250 Ala Tyr Phe Val Gly 255 Tyr Leu Lys Pro Thr 260 Thr Phe Met Leu	
aag tat gat gaa aat ggt aca atc aca gat gct gtt gat tgt tct caa	928
Lys Tyr Asp Glu Asn 270 Thr Ile Thr Asp Ala 275 Val Asp Cys Ser Gln 280	
aat cca ctt gct gaa ctc aaa tgc tct gtt aag agc ttt gag att gac	976
Asn Pro Leu Ala 285 Glu Leu Lys Cys Ser Val 290 Lys Ser Phe Glu 295 Ile Asp	
aaa gga att tac cag acc tct aat ttc agg gtt gtt ccc tca gga gat	1024
Lys Gly Ile Tyr 300 Gln Thr Ser Asn Phe 305 Arg Val Val Pro Ser 310 Gly Asp	
gtt gtg aga ttc cct aat att aca aac ttg tgt cct ttt gga gag gtt	1072
Val Val Arg 315 Phe Pro Asn Ile Thr 320 Asn Leu Cys Pro Phe 325 Gly Glu Val	
ttt aat gct act aaa ttc cct tct gtc tat gca tgg gag aga aaa aaa	1120
Phe Asn Ala Thr Lys Phe 335 Pro 335 Ser Val Tyr Ala Trp 340 Glu Arg Lys Lys	
att tct aat tgt gtt gct gat tac tct gtg ctc tac aac tca aca ttt	1168
Ile Ser Asn Cys Val 350 Ala Asp Tyr Ser Val 355 Leu Tyr Asn Ser Thr Phe 360	
ttt tca acc ttt aag tgc tat ggc gtt tct gcc act aag ttg aat gat	1216
Phe Ser Thr Phe 365 Lys Cys Tyr Gly Val Ser 370 Ala Thr Lys Leu Asn 375 Asp	
ctt tgc ttc tcc aat gtc tat gca gat tct ttt gta gtc aag gga gat	1264
Leu Cys Phe Ser 380 Asn Val Tyr Ala Asp 385 Ser Phe Val Val 390 Lys Gly Asp	
gat gta aga caa ata gcg cca gga caa act ggt gtt att gct gat tat	1312
Asp Val Arg 395 Gln Ile Ala Pro Gly 400 Gln Thr Gly Val 405 Ile Ala Asp Tyr	
aat tat aaa ttg cca gat gat ttc atg ggt tgt gtc ctt gct tgg aat	1360
Asn Tyr 410 Lys Leu Pro Asp 415 Phe Met Gly Cys Val 420 Leu Ala Trp Asn	
act agg aac att gat gct act tca act ggt aat tat aat tat aaa tat	1408
Thr Arg Asn Ile Asp Ala Thr Ser Thr Gly Asn 435 Tyr Asn Tyr Lys Tyr 440	
agg tat ctt aga cat ggc aag ctt agg ccc ttt gag aga gac ata tct	1456
Arg Tyr Leu Arg His 445 Gly Lys Leu Arg Pro 450 Phe Glu Arg Asp Ile Ser	
aat gtg cct ttc tcc cct gat ggc aaa cct tgc acc cca cct gct ctt	1504
Asn Val Pro Phe 460 Ser Pro Asp Gly Lys 465 Pro Cys Thr Pro Pro 470 Ala Leu	
aat tgt tat tgg cca tta aat gat tat ggt ttt tac acc act act ggc	1552
Asn Cys Tyr 475 Trp Pro Leu Asn Asp 480 Tyr Gly Phe Tyr Thr 485 Thr Thr Gly	

att Ile	ggc Gly 490	tac Tyr	caa Gln	cct Pro	tac Tyr	aga Arg 495	ggt Val	gta Val	gta Val	ctt Leu	tct Ser 500	ttt Phe	gaa Glu	ctt Leu	tta Leu	1600
aat Asn 505	gca Ala	ccg Pro	gcc Ala	acg Thr	ggt Val 510	tgt Cys	gga Gly	cca Pro	aaa Lys	tta Leu 515	tcc Ser	act Thr	gac Asp	ctt Leu	att Ile 520	1648
aag Lys	aac Asn	cag Gln	tgt Cys	gtc Val 525	aat Asn	ttt Phe	aat Asn	ttt Phe	aat Asn 530	gga Gly	ctc Leu	act Thr	ggg Gly	act Thr 535	ggg Gly	1696
gtg Val	tta Leu	act Thr	cct Pro 540	tct Ser	tca Ser	aag Lys	aga Arg	ttt Phe 545	caa Gln	cca Pro	ttt Phe	caa Gln 550	caa Gln	ttt Phe	ggc Gly	1744
cgt Arg	gat Asp	ggt Val 555	tct Ser	gat Asp	ttc Phe	act Thr 560	gat Asp	tcc Ser	ggt Val	cga Arg	gat Asp 565	cct Pro 565	aaa Lys	aca Thr	tct Ser	1792
gaa Glu 570	ata Ile	tta Leu	gac Asp	att Ile	tca Ser	cct Pro 575	tgc Cys	tct Ser	ttt Phe	ggg Gly	ggt Gly 580	gta Val	agt Ser	gta Val	att Ile	1840
aca Thr 585	cct Pro	gga Gly	aca Thr	aat Asn	gct Ala 590	tca Ser	tct Ser	gaa Glu	ggt Val	gct Ala 595	ggt Val	cta Leu	tat Tyr	caa Gln	gat Asp 600	1888
ggt Val	aac Asn	tgc Cys	act Thr	gat Asp 605	ggt Val	tct Ser	aca Thr	gca Ala	att Ile 610	cat His	gca Ala	gat Asp	caa Gln	ctc Leu 615	aca Thr	1936
cca Pro	gct Ala	tgg Trp	cgc Arg 620	ata Ile	tat Tyr	tct Ser	act Thr	gga Gly 625	aac Asn	aat Asn	gta Val	ttc Phe	cag Gln 630	act Thr	caa Gln	1984
gca Ala	ggc Gly	tgt Cys 635	ctt Leu	ata Ile	gga Gly	gct Ala	gag Glu 640	cat His	gtc Val	gac Asp	act Thr	tct Ser 645	tat Tyr	gag Glu	tgc Cys	2032
gac Asp	att Ile 650	cct Pro	att Ile	gga Gly	gct Ala	ggc Gly 655	att Ile	tgt Cys	gct Ala	agt Ser	tac Tyr 660	cat His	aca Thr	ggt Val	tct Ser	2080
tta Leu 665	tta Leu	cgt Arg	agt Ser	act Thr	agc Ser 670	caa Gln	aaa Lys	tct Ser	att Ile	gtg Val 675	gct Ala	tat Tyr	act Thr	atg Met	tct Ser 680	2128
tta Leu	ggg Gly	gct Ala	gat Asp	agt Ser 685	tca Ser	att Ile	gct Ala	tac Tyr	tct Ser 690	aat Asn	aac Asn	acc Thr	att Ile	gct Ala 695	ata Ile	2176
cct Pro	act Thr	aac Asn	ttt Phe 700	tca Ser	att Ile	agc Ser	att Ile	act Thr 705	aca Thr	gaa Glu	gta Val	atg Met	cct Pro 710	ggt Val	tct Ser	2224
atg Met	gct Ala	aaa Lys 715	acc Thr	tcc Ser	gta Val	gat Asp	tgt Cys 720	aat Asn	atg Met	tac Tyr	atc Ile	tgc Cys 725	gga Gly	gat Asp	tct Ser	2272
act Thr 730	gaa Glu	tgt Cys	gct Ala	aat Asn	ttg Leu	ctt Leu 735	ctc Leu	caa Gln	tat Tyr	ggg Gly	agc Ser 740	ttt Phe	tgc Cys	aca Thr	caa Gln	2320
cta Leu 745	aat Asn	cgt Arg	gca Ala	ctc Leu	tca Ser 750	ggg Gly	att Ile	gct Ala	gct Ala	gaa Glu 755	cag Gln	gat Asp	cgc Arg	aac Asn	aca Thr 760	2368

cgt Arg	gaa Glu	gtg Val	ttc Phe	gct Ala 765	caa Gln	gtc Val	aaa Lys	caa Gln	atg Met 770	tac Tyr	aaa Lys	acc Thr	cca Pro	act Thr 775	ttg Leu	2416
aaa Lys	tat Tyr	ttt Phe	ggt Gly 780	ggt Gly	ttt Phe	aat Asn	ttt Phe	tca Ser 785	caa Gln	ata Ile	tta Leu	cct Pro	gac Asp 790	cct Pro	cta Leu	2464
aag Lys	cca Pro	act Thr 795	aag Lys	agg Arg	tct Ser	ttt Phe	att Ile 800	gag Glu	gac Asp	ttg Leu	ctc Leu	ttt Phe 805	aat Asn	aag Lys	gtg Val	2512
aca Thr	ctc Leu 810	gct Ala	gat Asp	gct Ala	ggc Gly	ttc Phe 815	atg Met	aag Lys	caa Gln	tat Tyr	ggc Gly 820	gaa Glu	tgc Cys	cta Leu	ggt Gly	2560
gat Asp 825	att Ile	aat Asn	gct Ala	aga Arg	gat Asp 830	ctc Leu	att Ile	tgt Cys	gcg Ala	cag Gln 835	aag Lys	ttc Phe	aat Asn	gga Gly	ctt Leu 840	2608
aca Thr	gtg Val	ttg Leu	cca Pro	cct Pro 845	ctg Leu	ctc Leu	act Thr	gat Asp	gat Asp 850	atg Met	att Ile	gct Ala	gcc Ala	tac Tyr 855	act Thr	2656
gct Ala	gct Ala	cta Leu	gtt Val 860	agt Ser	ggc Gly	act Thr	gcc Ala	act Thr 865	gct Ala	gga Gly	tgg Trp	aca Thr	ttt Phe 870	ggc Gly	gct Ala	2704
ggc Gly	gct Ala	gct Ala 875	ctt Leu	caa Gln	ata Ile	cct Pro	ttt Phe 880	gct Ala	atg Met	caa Gln	atg Met 885	gca Ala	tat Tyr	agg Arg	ttc Phe	2752
aat Asn 890	ggc Gly	att Ile	gga Gly	gtt Val	acc Thr	caa Gln 895	aat Asn	gtt Val	ctc Leu	tat Tyr	gag Glu 900	aac Asn	caa Gln	aaa Lys	caa Gln	2800
atc Ile 905	gcc Ala	aac Asn	caa Gln	ttt Phe	aac Asn 910	aag Lys	gcg Ala	att Ile	agt Ser	caa Gln 915	att Ile	caa Gln	gaa Glu	tca Ser	ctt Leu 920	2848
aca Thr	aca Thr	aca Thr	tca Ser	act Thr 925	gca Ala	ttg Leu	ggc Gly	aag Lys	ctg Leu 930	caa Gln	gac Asp	gtt Val	gtt Val	aac Asn 935	cag Gln	2896
aat Asn	gct Ala	caa Gln	gca Ala 940	tta Leu	aac Asn	aca Thr	ctt Leu	gtt Val 945	aaa Lys	caa Gln	ctt Leu	agc Ser	tct Ser 950	aat Asn	ttt Phe	2944
ggc Gly	gca Ala	att Ile 955	tca Ser	agt Ser	gtg Val	cta Leu	aat Asn 960	gat Asp	atc Ile	ctt Leu	tcg Ser	cga Arg 965	ctt Leu	gat Asp	aaa Lys	2992
gtc Val	gag Glu 970	gcg Ala	gag Glu	gta Val	caa Gln	att Ile 975	gac Asp	agg Arg	tta Leu	att Ile	aca Thr 980	ggc Gly	aga Arg	ctt Leu	caa Gln	3040
agc Ser 985	ctt Leu	caa Gln	acc Thr	tat Tyr	gta Val 990	aca Thr	caa Gln	caa Gln	cta Leu	atc Ile 995	agg Arg	gct Ala	gct Ala	gaa Glu	atc Ile 1000	3088
agg Arg	gct Ala	tct Ser	gct Ala	aat Asn 1005	ctt Leu	gct Ala	gct Ala	act Thr	aaa Lys 1010	atg Met	tct Ser	gag Glu	tgt Cys	gtt Val 1015		3133
ctt Leu	gga Gly	caa Gln	tca Ser	aaa Lys 1020	aga Arg	gtt Val	gac Asp	ttt Phe	tgt Cys 1025	gga Gly	aag Lys	ggc Gly	tac Tyr	cac His 1030		3178

ctt atg tcc ttc cca Leu Met Ser Phe Pro 1035	caa gca gcc ccg cat Gln Ala Ala Pro His 1040	ggt gtt gtc ttc cta Gly Val Val Phe Leu 1045	3223
cat gtc acg tat gtg His Val Thr Tyr Val 1050	cca tcc cag gag agg Pro Ser Gln Glu Arg 1055	aac ttc acc aca gcg Asn Phe Thr Thr Ala 1060	3268
cca gca att tgt cat Pro Ala Ile Cys His 1065	gaa ggc aaa gca tac Glu Gly Lys Ala Tyr 1070	ttc cct cgt gaa ggt Phe Pro Arg Glu Gly 1075	3313
gtt ttt gtg ttt aat Val Phe Val Phe Asn 1080	ggc act tct tgg ttt Gly Thr Ser Trp Phe 1085	att aca cag agg aac Ile Thr Gln Arg Asn 1090	3358
ttc ttt tct cca caa Phe Phe Ser Pro Gln 1095	ata att act aca gac Ile Ile Thr Thr Asp 1100	aac aca ttt gtc tca Asn Thr Phe Val Ser 1105	3403
gga aat tgt gat gtc Gly Asn Cys Asp Val 1110	gtt att ggc atc att Val Ile Gly Ile Ile 1115	aac aac aca gtt tat Asn Asn Thr Val Tyr 1120	3448
gat cct ctg caa cct Asp Pro Leu Gln Pro 1125	gag ctt gac tca ttc Glu Leu Asp Ser Phe 1130	aaa gaa gag ctg gac Lys Glu Glu Leu Asp 1135	3493
aag tac ttc aaa aat Lys Tyr Phe Lys Asn 1140	cat aca tca cca gat His Thr Ser Pro Asp 1145	gtt gat ctt ggc gac Val Asp Leu Gly Asp 1150	3538
att tca ggc att aac Ile Ser Gly Ile Asn 1155	gct tct gtc gtc aac Ala Ser Val Val Asn 1160	att caa aaa gaa att Ile Gln Lys Glu Ile 1165	3583
gac cgc ctc aat gag Asp Arg Leu Asn Glu 1170	gtc gct aaa aat tta Val Ala Lys Asn Leu 1175	aat gaa tca ctc att Asn Glu Ser Leu Ile 1180	3628
gac ctt caa gaa ttg Asp Leu Gln Glu Leu 1185	gga aaa tat gag caa Gly Lys Tyr Glu Gln 1190	tat att aaa tgg cct Tyr Ile Lys Trp Pro 1195	3673
tgg tat gtt tgg ctc Trp Tyr Val Trp Leu 1200	ggc ttc att gct gga Gly Phe Ile Ala Gly 1205	cta att gcc atc gtc Leu Ile Ala Ile Val 1210	3718
atg gtt aca atc ttg Met Val Thr Ile Leu 1215	ctt tgt tgc atg act Leu Cys Cys Met Thr 1220	agt tgt tgc agt tgc Ser Cys Cys Ser Cys 1225	3763
ctc aag ggt gca tgc Leu Lys Gly Ala Cys 1230	tct tgt ggt tct tgc Ser Cys Gly Ser Cys 1235	tgc aag ttt gat gag Cys Lys Phe Asp Glu 1240	3808
gat gac tct gag cca Asp Asp Ser Glu Pro 1245	gtt ctc aag ggt gtc Val Leu Lys Gly Val 1250	aaa tta cat tac aca Lys Leu His Tyr Thr 1255	3853
taaacgaact tatggatttg tttatgagat tttttactct tggatcaatt actgcacagc			3913
cagtaaaaaat tgacaatgct tctcctgcaa gt			3945

<210> 3
<211> 1255
<212> PRT

ES 2 396 127 T3

<213> CORONAVIRUS

<400> 3

```

Met Phe Ile Phe Leu Leu Phe Leu Thr Leu Thr Ser Gly Ser Asp Leu
 1      5      10      15

Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
 20      25      30

His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
 35      40      45

Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
 50      55      60

Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
 65      70      75      80

Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
 85      90      95

Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
100      105      110

Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys
115      120      125

Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met
130      135      140

Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr
145      150      155      160

Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser
165      170      175

Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly
180      185      190

Phe Leu Tyr Val Tyr Lys Gly Tyr Gln Pro Ile Asp Val Val Arg Asp
195      200      205

Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu
210      215      220

Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro
225      230      235      240

```

Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala Tyr Phe Val Gly Tyr
245 250 255

Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp Glu Asn Gly Thr Ile
260 265 270

Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys
275 280 285

Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser Asn
290 295 300

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

Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser
325 330 335

Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr
340 345 350

Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly
355 360 365

Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala
370 375 380

Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala Pro Gly
385 390 395 400

Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe
405 410 415

Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser
420 425 430

Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu
435 440 445

Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro Phe Ser Pro Asp Gly
450 455 460

Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp
465 470 475 480

Tyr Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val
485 490 495

Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly
500 505 510

Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn
 515 520 525
 Phe Asn Gly Leu Thr Gly Thr Gly Val Leu Thr Pro Ser Ser Lys Arg
 530 535 540
 Phe Gln Pro Phe Gln Gln Phe Gly Arg Asp Val Ser Asp Phe Thr Asp
 545 550 555 560
 Ser Val Arg Asp Pro Lys Thr Ser Glu Ile Leu Asp Ile Ser Pro Cys
 565 570 575
 Ser Phe Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Ala Ser Ser
 580 585 590
 Glu Val Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Asp Val Ser Thr
 595 600 605
 Ala Ile His Ala Asp Gln Leu Thr Pro Ala Trp Arg Ile Tyr Ser Thr
 610 615 620
 Gly Asn Asn Val Phe Gln Thr Gln Ala Gly Cys Leu Ile Gly Ala Glu
 625 630 635 640
 His Val Asp Thr Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile
 645 650 655
 Cys Ala Ser Tyr His Thr Val Ser Leu Leu Arg Ser Thr Ser Gln Lys
 660 665 670
 Ser Ile Val Ala Tyr Thr Met Ser Leu Gly Ala Asp Ser Ser Ile Ala
 675 680 685
 Tyr Ser Asn Asn Thr Ile Ala Ile Pro Thr Asn Phe Ser Ile Ser Ile
 690 695 700
 Thr Thr Glu Val Met Pro Val Ser Met Ala Lys Thr Ser Val Asp Cys
 705 710 715 720
 Asn Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ala Asn Leu Leu Leu
 725 730 735
 Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala Leu Ser Gly Ile
 740 745 750
 Ala Ala Glu Gln Asp Arg Asn Thr Arg Glu Val Phe Ala Gln Val Lys
 755 760 765
 Gln Met Tyr Lys Thr Pro Thr Leu Lys Tyr Phe Gly Gly Phe Asn Phe
 770 775 780

ES 2 396 127 T3

Ser Gln Ile Leu Pro Asp Pro Leu Lys Pro Thr Lys Arg Ser Phe Ile
785 790 795 800

Glu Asp Leu Leu Phe Asn Lys Val Thr Leu Ala Asp Ala Gly Phe Met
805 810 815

Lys Gln Tyr Gly Glu Cys Leu Gly Asp Ile Asn Ala Arg Asp Leu Ile
820 825 830

Cys Ala Gln Lys Phe Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr
835 840 845

Asp Asp Met Ile Ala Ala Tyr Thr Ala Ala Leu Val Ser Gly Thr Ala
850 855 860

Thr Ala Gly Trp Thr Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe
865 870 875 880

Ala Met Gln Met Ala Tyr Arg Phe Asn Gly Ile Gly Val Thr Gln Asn
885 890 895

Val Leu Tyr Glu Asn Gln Lys Gln Ile Ala Asn Gln Phe Asn Lys Ala
900 905 910

Ile Ser Gln Ile Gln Glu Ser Leu Thr Thr Thr Ser Thr Ala Leu Gly
915 920 925

Lys Leu Gln Asp Val Val Asn Gln Asn Ala Gln Ala Leu Asn Thr Leu
930 935 940

Val Lys Gln Leu Ser Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn
945 950 955 960

Asp Ile Leu Ser Arg Leu Asp Lys Val Glu Ala Glu Val Gln Ile Asp
965 970 975

Arg Leu Ile Thr Gly Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln
980 985 990

Gln Leu Ile Arg Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu Ala Ala
995 1000 1005

Thr Lys Met Ser Glu Cys Val Leu Gly Gln Ser Lys Arg Val Asp
1010 1015 1020

Phe Cys Gly Lys Gly Tyr His Leu Met Ser Phe Pro Gln Ala Ala
1025 1030 1035

Pro His Gly Val Val Phe Leu His Val Thr Tyr Val Pro Ser Gln
1040 1045 1050

Glu Arg Asn Phe Thr Thr Ala Pro Ala Ile Cys His Glu Gly Lys
 1055 1060 1065
 Ala Tyr Phe Pro Arg Glu Gly Val Phe Val Phe Asn Gly Thr Ser
 1070 1075 1080
 Trp Phe Ile Thr Gln Arg Asn Phe Phe Ser Pro Gln Ile Ile Thr
 1085 1090 1095
 Thr Asp Asn Thr Phe Val Ser Gly Asn Cys Asp Val Val Ile Gly
 1100 1105 1110
 Ile Ile Asn Asn Thr Val Tyr Asp Pro Leu Gln Pro Glu Leu Asp
 1115 1120 1125
 Ser Phe Lys Glu Glu Leu Asp Lys Tyr Phe Lys Asn His Thr Ser
 1130 1135 1140
 Pro Asp Val Asp Leu Gly Asp Ile Ser Gly Ile Asn Ala Ser Val
 1145 1150 1155
 Val Asn Ile Gln Lys Glu Ile Asp Arg Leu Asn Glu Val Ala Lys
 1160 1165 1170
 Asn Leu Asn Glu Ser Leu Ile Asp Leu Gln Glu Leu Gly Lys Tyr
 1175 1180 1185
 Glu Gln Tyr Ile Lys Trp Pro Trp Tyr Val Trp Leu Gly Phe Ile
 1190 1195 1200
 Ala Gly Leu Ile Ala Ile Val Met Val Thr Ile Leu Leu Cys Cys
 1205 1210 1215
 Met Thr Ser Cys Cys Ser Cys Leu Lys Gly Ala Cys Ser Cys Gly
 1220 1225 1230
 Ser Cys Cys Lys Phe Asp Glu Asp Asp Ser Glu Pro Val Leu Lys
 1235 1240 1245
 Gly Val Lys Leu His Tyr Thr
 1250 1255

<210> 4
 <211> 3943
 <212> ADN
 <213> CORONAVIRUS
 <400> 4

ctcttctgga	aaaaggtagg	cttatcatta	gagaaaacaa	cagagttgtg	gtttcaagtg	60
atattcttgt	taacaactaa	acgaacatgt	ttattttctt	attatttctt	actctcacta	120
gtggtagtga	ccttgaccgg	tgcaccactt	ttgatgatgt	tcaagctcct	aattacactc	180
aacatacttc	atctatgagg	ggggtttact	atcctgatga	aatTTTTaga	tcagacactc	240
tttatttaac	tcaggattta	tttcttccat	tttattctaa	tgttacaggg	tttcatacta	300
ttaatcatac	gtttggcaac	cctgtcatac	cttttaagga	tggatatttat	tttgctgcc	360
cagagaaatc	aaatgtttgtc	cgtgggttggg	tttttggttc	taccatgaac	aacaagtcac	420
agtcggtgat	tattattaac	aattctacta	atgttggttat	acgagcatgt	aactttgaat	480
tgtgtgacaa	ccctttcttt	gctgtttcta	aacccatggg	tacacagaca	catactatga	540
tattcgataa	tgcatttaac	tgcactttcg	agtacatatc	tgatgccttt	tcgcttgatg	600
tttcagaaaa	gtcaggtaac	tttaaacact	tacgagagtt	tgtgtttaaa	aataaagatg	660
ggtttctcta	tgtttataag	ggctatcaac	ctatagatgt	agttcgtgat	ctaccttctg	720
gttttaacac	tttgaaacct	atttttaagt	tgccctttgg	tattaacatt	acaaatttta	780
gagccattct	tacagccttt	tcacctgttc	aagacatttg	gggcacgtca	gctgcagcct	840
attttggttg	ctatttaaag	ccaactacat	ttatgctcaa	gtatgatgaa	aatgggtacaa	900
tcacagatgc	tgttgattgt	tctcaaaatc	cacttgctga	actcaaatgc	tctgttaaga	960
gctttgagat	tgacaaagga	atttaccaga	cctctaattt	cagggttggt	ccctcaggag	1020
atgttggtgag	attccctaac	attacaaact	tgtgtccttt	tggagagggt	tttaatgcta	1080
ctaaattccc	ttctgtctat	gcatgggaga	gaaaaaaaaat	ttctaattgt	gttgctgatt	1140
actctgtgct	ctacaactca	acattttttt	caacctttta	gtgctatggc	gtttctgcc	1200
ctaagttgaa	tgatctttgc	ttctccaatg	tctatgcaga	ttcttttgta	gtcaagggag	1260
atgatgtaag	acaaatagcg	ccaggacaaa	ctgggtgttat	tgctgattat	aattataaat	1320
tgccagatga	ttcatgggt	tgtgtccttg	cttggaatac	taggaacatt	gatgctactt	1380
caactggtaa	ttataattat	aaatataggt	atcttagaca	tggcaagctt	aggccctttg	1440
agagagacat	atctaattgt	cctttctccc	ctgatggcaa	accttgcacc	ccacctgttc	1500
ttaattgita	ttggccatta	aatgattatg	gtttttacac	cactactggc	attggctacc	1560
aaccttacag	agttgtagta	ctttcttttg	aacttttaaa	tgcaccggcc	acggtttgtg	1620
gacaaaaatt	atccactgac	cttattaaga	accagtgtgt	caattttaat	tttaatggac	1680
tcactggtac	tgggtgtgta	actccttctt	caaagagatt	tcaaccattt	caacaatttg	1740
gccgtgatgt	ctctgatttc	actgattccg	ttcgagatcc	taaaacatct	gaaatattag	1800
acatttcacc	ttgctctttt	gggggtgtaa	gtgtaattac	acctggaaca	aatgcttcat	1860
ctgaagttgc	tgttctatat	caagatgtta	actgcactga	tgtttctaca	gcaatccatg	1920
cagatcaact	cacaccagct	tggcgcatat	attctactgg	aaacaatgta	ttccagactc	1980
aagcaggctg	tcttatagga	gctgagcatg	tcgacacttc	ttatgagtgc	gacatttcta	2040

ttggagctgg	catttgtgct	agttaccata	cagtttcttt	attacgtagt	actagccaaa	2100
aatctattgt	ggcttatact	atgtcttttag	gtgctgatag	ttcaattgct	tactctaata	2160
acaccattgc	tatacctact	aacttttcaa	ttagcattac	tacagaagta	atgcctgttt	2220
ctatggctaa	aacctccgta	gattgtaata	tgtacatctg	cggagattct	actgaatgtg	2280
ctaatttgct	tctccaatat	ggtagctttt	gcacacaact	aaatcgtgca	ctctcaggta	2340
ttgctgctga	acaggatcgc	aacacacgtg	aagtgttcgc	tcaagtcaaa	caaatgtaca	2400
aaacccaac	tttgaaatat	tttggtggtt	ttaatttttc	acaaatatta	cctgaccctc	2460
taaagccaac	taagaggtct	tttattgagg	acttgctctt	taataagggtg	acactcgctg	2520
atgctggctt	catgaagcaa	tatggcgaat	gcctaggtga	tattaatgct	agagatctca	2580
tttgtgcgca	gaagttcaat	gggcttacag	tgttgccacc	tctgctcact	gatgatatga	2640
ttgctgccta	cactgctgct	ctagtttagtg	gtactgccac	tgctggatgg	acatttggtg	2700
ctggcgctgc	tcttcaaata	ccttttgcta	tgcaaatggc	atatagggttc	aatggcattg	2760
gagttacca	aaatgttctc	tatgagaacc	aaaaacaaat	cgccaaccaa	tttaacaagg	2820
cgattagtca	aattcaagaa	tcacttacaa	caacatcaac	tgcatggggc	aagctgcaag	2880
acgttggttaa	ccagaatgct	caagcattaa	acacacttgt	taaacaactt	agctctaatt	2940
ttggtgcaat	ttcaagtgtg	ctaaatgata	tcctttcgcg	acttgataaa	gtcgaggcgg	3000
aggtacaaat	tgacaggcta	attacaggca	gacttcaaag	ccttcaaacc	tatgtaacac	3060
aacaactaat	cagggctgct	gaaatcaggg	cttctgctaa	tcttgctgct	actaaaatgt	3120
ctgagtgtgt	tcttggaaca	tcaaaaagag	ttgacttttg	tggaaagggc	taccacctta	3180
tgtccttccc	acaagcagcc	ccgcatgggtg	ttgtcttcct	acatgtcacg	tatgtgccat	3240
cccaggagag	gaacttcacc	acagcgccag	caatttgctca	tgaaggcaaa	gcatacttcc	3300
ctcgtgaagg	tgtttttgtg	tttaatggca	cttcttggtt	tattacacag	aggaacttct	3360
tttctccaca	aataattact	acagacaata	catttgtctc	aggaaattgt	gatgtcgtta	3420
ttggcatcat	taacaacaca	gtttatgata	ctctgcaacc	tgagcttgac	tcattcaaag	3480
aagagctgga	caagtacttc	aaaaatcata	catcaccaga	tggtgatctt	ggcgacattt	3540
caggcattaa	cgcttctgtc	gtcaacattc	aaaaagaaat	tgaccgcctc	aatgagggtcg	3600
ctaaaaattt	aatgaatca	ctcattgacc	ttcaagaatt	gggaaaatat	gagcaatata	3660
ttaaatggcc	ttggtatggt	tggctcggct	tcattgctgg	actaattgcc	atcgtcatgg	3720
ttacaatctt	gctttgttgc	atgactagtt	gttgcaagg	cctcaagggg	gcatgctctt	3780
gtggttcttg	ctgcaagttt	gatgaggatg	actctgagcc	agttctcaag	ggtgtcaaat	3840
tacattacac	ataaacgaac	ttatggattt	gtttatgaga	ttttttactc	ttggatcaat	3900
tactgcacag	ccagtaaaaa	ttgacaatgc	ttctcctgca	agt		3943

<210> 5

<211> 2049

<212> ADN

<213> CORONAVIRUS

<400> 5

```

ctcttctgga aaaaggtagg cttatcatta gagaaaacaa cagagttgtg gtttcaagtg      60
atattcttgt taacaactaa acgaacatgt ttattttctt attatttctt actctcacta     120
gtggtagtga ccttgaccgg tgcaccactt ttgatgatgt tcaagctcct aattacactc     180
aacatacttc atctatgagg ggggtttact atcctgatga aattttttaga tcagacactc     240
tttatttaac tcaggattta tttcttccat ttatttctaa tgttacaggg tttcatacta     300
ttaatcatac gtttggaac cctgtcatac cttttaagga tgggtatttat tttgctgccca     360
cagagaaatc aaatgttgtc cgtggttggg tttttggttc taccatgaac aacaagtcac     420
agtcggtgat tattattaac aattctacta atgttggttat acgagcatgt aactttgaat     480
tgtgtgacaa ccccttcttt gctgtttcta aacccatggg tacacagaca catactatga     540
tattcgataa tgcatttaac tgcactttcg agtacatata tgatgccttt tcgcttgatg     600
tttcagaaaa gtcaggtaac tttaaacact tacgagagtt tgtgtttaaa aataaagatg     660
ggtttctcta tgtttataag ggctatcaac ctatagatgt agttcgtgat ctaccttctg     720
gttttaacac ttgaaacct atttttaagt tgcctcttgg tattaacatt acaaatttta     780
gagccattct tacagccttt tcacctgctc aagacatttg gggcacgtca gctgcagcct     840
attttgttgg ctatttaaa ccaactacat ttatgctcaa gtatgatgaa aatggtacaa     900
tcacagatgc tgttgattgt tctcaaaatc cacttgctga actcaaatgc tctgttaaga     960
gctttgagat tgacaaagga atttaccaga cctctaattt cagggttggt ccctcaggag    1020
atgttgtagg attccctaatt attacaaact tgtgtccttt tggagagggt tttaatgcta    1080
ctaaattccc ttctgtctat gcatgggaga gaaaaaaaaat ttctaattgt gttgctgatt    1140
actctgtgct ctacaactca acatTTTTTT caacctttaa gtgctatggc gtttctgccca    1200
ctaagttgaa tgatctttgc ttctccaatg tctatgcaga ttcttttgta gtcaagggag    1260
atgatgtaag acaaatagcg ccaggacaaa ctggtgttat tgctgattat aattataaat    1320
tgccagatga tttcatgggt tgtgtccttg cttggaatac taggaacatt gatgctactt    1380
caactggtaa ttataattat aaatataggt atcttagaca tggcaagctt aggccctttg    1440
agagagacat atctaattgt ctttctctcc ctgatggcaa accttgcacc ccacctgctc    1500
ttaattgtta ttggccatta aatgattatg gtttttacac cactactggc attggctacc    1560
aaccttacag agttgtagta ctttcttttg aacttttaaa tgcaccggcc acggtttgtg    1620
gaccaaaaatt atccactgac cttattaaga accagtgtgt caattttaat tttaatggac    1680
tcactggtag tggtgtgtta actccttctt caaagagatt tcaaccattt caacaatttg    1740
gccgtgatgt ctctgatttc actgattccg ttcgagatcc taaaacatct gaaatattag    1800

```

acatttcacc	ttgctctttt	gggggtgtaa	gtgtaattac	acctggaaca	aatgcttcat	1860
ctgaagttgc	tgttctatat	caagatgtta	actgcactga	tgtttctaca	gcaatccatg	1920
cagatcaact	cacaccagct	tggcgcatat	attctactgg	aaacaatgta	ttccagactc	1980
aagcaggctg	tcttatagga	gctgagcatg	tcgacacttc	ttatgagtg	gacattccta	2040
ttggagctg						2049

<210> 6
 <211> 2027
 <212> ADN
 <213> CORONAVIRUS

<400> 6

catgcagatc	aactcacacc	agcttggcgc	atatattcta	ctggaacaa	tgtattccag	60
actcaagcag	gctgtcttat	aggagctgag	catgtcgaca	cttcttatga	gtgcgacatt	120
cctattggag	ctggcatttg	tgctagttag	catacagttt	ctttattacg	tagtactagc	180
caaaaatcta	ttgtggctta	tactatgtct	ttaggtgctg	atagttcaat	tgcttactct	240
aataacacca	ttgctatacc	tactaacttt	tcaattagca	ttactacaga	agtaatgcct	300
gtttctatgg	ctaaaacctc	cgtagattgt	aatatgtaca	tctgcggaga	ttctactgaa	360
tgtgctaatt	tgcttctcca	atatggtagc	ttttgcacac	aactaaatcg	tgcaactctca	420
ggatttgctg	ctgaacagga	tcgcaacaca	cgtgaagtgt	tcgctcaagt	caaacaaatg	480
tacaaaaccc	caactttgaa	atattttggg	ggttttaatt	tttcacaaat	attacctgac	540
cctctaaagc	caactaagag	gtcttttatt	gaggacttgc	tctttaataa	ggtagacttc	600
gctgatgctg	gcttcatgaa	gcaatatggc	gaatgcctag	gtgatattaa	tgctagagat	660
ctcatttgctg	cgcagaagtt	caatgggctt	acagtgttgc	cacctctgct	cactgatgat	720
atgattgctg	cctacactgc	tgctctagtt	agtggtagctg	ccactgctgg	atggacattt	780
ggtagctggcg	ctgctcttca	aatacctttt	gctatgcaaa	tggcatatag	gttcaatggc	840
attggaggtta	cccaaaatgt	tctctatgag	aacaaaaaac	aaatcgccaa	ccaatttaac	900
aaggcgatta	gtcaaattca	agaatcactt	acaacaacat	caactgcatt	gggcaagctg	960
caagacgttg	ttaaccagaa	tgctcaagca	ttaaaccacac	ttgttaaaca	acttagctct	1020
aattttgggtg	caatttcaag	tgtgctaaat	gatatccttt	cgcgacttga	taaagtcgag	1080
gcggagggtac	aaattgacag	gttaattaca	ggcagacttc	aaagccttca	aacctatgta	1140
acacaacaac	taatcagggc	tgctgaaatc	agggcttctg	ctaactcttg	tgctactaaa	1200
atgtctgagt	gtgttcttgg	acaatcaaaa	agagttgact	tttgtggaaa	gggctaccac	1260
cttatgtcct	tcccacaagc	agccccgcat	gggtgtgtct	tcctacatgt	cacgtatgtg	1320
ccatcccagg	agaggaactt	caccacagcg	ccagcaattt	gtcatgaagg	caaagcatac	1380

ttccctcgtg	aaggtgtttt	tgtgtttaat	ggcacttctt	ggtttattac	acagaggaac	1440
ttcttttctc	cacaaataat	tactacagac	aatacatttg	tctcaggaaa	ttgtgatgtc	1500
gttattggcg	tcattaacaa	cacagtttat	gacctctgc	aacctgagct	tgactcattc	1560
aaagaagagc	tggacaagta	cttcaaaaat	catacatcac	cagatggtga	tcttggcgac	1620
atttcaggca	ttaacgcttc	tgtcgtcaac	attcaaaaag	aaattgaccg	cctcaatgag	1680
gtcgctaaaa	atttaaatga	atcactcatt	gaccttcaag	aattgggaaa	atatgagcaa	1740
tatattaaat	ggccttggtg	tgtttggctc	ggcttcattg	ctggactaat	tgccatcgtc	1800
atggttacia	tcttgctttg	ttgcatgact	agttgttgca	gttgccctca	gggtgcatgc	1860
tcttgtgggt	cttgctgcaa	gtttgatgag	gatgactctg	agccagttct	caaggggtgtc	1920
aaattacatt	acacataaac	gaacttatgg	atttgtttat	gagatTTTTT	actcttggat	1980
caattactgc	acagccagta	aaaattgaca	atgcttctcc	tgcaagt		2027

<210> 7

<211> 1096

<212> ADN

<213> CORONAVIRUS

<400> 7

tcttgctttg	ttgcatgact	agttgttgca	gttgccctca	gggtgcatgc	tcttgtgggt	60
cttgctgcaa	gtttgatgag	gatgactctg	agccagttct	caaggggtgtc	aaattacatt	120
acacataaac	gaacttatgg	atttgtttat	gagatTTTTT	actcttggat	caattactgc	180
acagccagta	aaaattgaca	atgcttctcc	tgcaagtact	gttcatgcta	cagcaacgat	240
accgctacia	gcctcactcc	ctttcggatg	gcttgttatt	ggcgttgcat	ttcttgctgt	300
ttttcagagc	gctacaaaaa	taattgcgct	caataaaaga	tggcagctag	ccctttataa	360
gggcttccag	ttcatttgca	atttactgct	gctatttggt	accatctatt	cacatctttt	420
gcttgtcgct	gcagggtatg	aggcgcaatt	tttgtacctc	tatgccttga	tatattttct	480
acaatgcctc	aacgcctgta	gaattattat	gagatgttgg	ctttgttgga	agtgc aaatc	540
caagaacca	ttacttttat	atgccaaacta	ctttgtttgc	tggcacacac	ataactatga	600
ctactgtata	ccatataaca	gtgtcacaga	tacaattgtc	gttactgaag	gtgacggcat	660
ttcaacacca	aaactcaaag	aagactacca	aattgggtgg	tattctgagg	ataggcactc	720
aggtgttaaa	gactatgtcg	ttgtacatgg	ctatttcacc	gaagtttact	accagcttga	780
gtctacacia	attactacag	acactgggtat	tgaaaatgct	acattcttca	tctttaacia	840
gcttgttaaa	gacccaccga	atgtgcaaat	acacacaatc	gacggctctt	caggagttgc	900
taatccagca	atggatccaa	tttatgatga	gccgacgacg	actactagcg	tgccctttgta	960
agcacaagaa	agtgagtacg	aacttatgta	ctcattcggt	tcggaagaaa	caggtacgtt	1020
aatagttaat	agcgtacttc	tttttcttgc	tttcgtggta	ttcttgctag	tcacactagc	1080
catccttact	gcgctt					1096

ES 2 396 127 T3

<210> 8
 <211> 1135
 <212> ADN
 <213> CORONAVIRUS

<400> 8

```

attgccatcg tcatggttac aatcttgctt tgttgcata ctagttgttg cagttgcctc      60
aaggggtgcat gctcttggtg ttcttgctgc aagtttgatg aggatgactc tgagccagtt    120
ctcaaggggtg tcaaattaca ttacacataa acgaacttat ggatttggtt atgagatgtt    180
ttactcttggt atcaattact gcacagccag taaaaattga caatgcttct cctgcaagta    240
ctgttcatgc tacagcaacg ataccgctac aagcctcact ccctttcgga tggccttgta    300
ttggcggttg atttcttgct gtttttcaga gcgctaccaa aataattgct ctcaataaaa    360
gatggcagct agccctttat aagggcttcc agttcatttg caatttactg ctgctatttg    420
ttaccatcta ttcacatctt ttgcttgctg ctgcaggtat ggaggcgcaa tttttgtacc    480
tctatgcctt gatataatgt ctacaatgca tcaacgcata tagaattatt atgagatgtt    540
ggctttgttg gaagtgcata tccaagaacc cattacttta tgatgccaac tactttgttt    600
gctggcacac acataactat gactactgta taccatataa cagtgtcaca gatacaattg    660
tcgttactga aggtgacggc atttcaacac caaaactcaa agaagactac caaattggtg    720
gttattctga ggataggcac tcagggtgta aagactatgt cgttgatcat ggctattttca    780
ccgaagttaa ctaccagctt gagtctacac aaattactac agacactggt attgaaaatg    840
ctacattctt catctttaac aagcttggtt aagaccacc gaatgtgcaa atacacacaa    900
tcgacggctc ttcaggagtt gctaattccag caatggatcc aatttatgat gagccgacga    960
cgactactag cgtgcctttg taagcacaag aaagtgagta cgaacttatg tactcattcg   1020
tttcggaaga aacaggtacg ttaatagtta atagcgtact tctttttctt gctttcgtgg   1080
tattcttgct agtcacacta gccatcctta ctgcgcttcg attgtgtgct tactg      1135

```

<210> 9
 <211> 1096
 <212> ADN
 <213> CORONAVIRUS

<220>
 <221> CDS
 <222> (137)..(958)
 <223>

<400> 9

tcttgctttg ttgcatgact agttgttgca gttgcctcaa ggggtgcatgc tcttgtgggtt	60
cttgctgcaa gtttgatgag gatgactctg agccagttct caagggtgtc aaattacatt	120
acacataaac gaactt atg gat ttg ttt atg aga ttt ttt act ctt gga tca	172
Met Asp Leu Phe Met Arg Phe Phe Thr Leu Gly Ser	
1 5 10	
att act gca cag cca gta aaa att gac aat gct tct cct gca agt act	220
Ile Thr Ala Gln Pro Val Lys Ile Asp Asn Ala Ser Pro Ala Ser Thr	
15 20 25	
gtt cat gct aca gca acg ata ccg cta caa gcc tca ctc cct ttc gga	268
Val His Ala Thr Ala Thr Ile Pro Leu Gln Ala Ser Leu Pro Phe Gly	
30 35 40	
tgg ctt gtt att ggc gtt gca ttt ctt gct gtt ttt cag agc gct acc	316
Trp Leu Val Ile Gly Val Ala Phe Leu Ala Val Phe Gln Ser Ala Thr	
45 50 55 60	
aaa ata att gcg ctc aat aaa aga tgg cag cta gcc ctt tat aag ggc	364
Lys Ile Ile Ala Leu Asn Lys Arg Trp Gln Leu Ala Leu Tyr Lys Gly	
65 70 75	
ttc cag ttc att tgc aat tta ctg ctg cta ttt gtt acc atc tat tca	412
Phe Gln Phe Ile Cys Asn Leu Leu Leu Phe Val Thr Ile Tyr Ser	
80 85 90	
cat ctt ttg ctt gtc gct gca ggt atg gag gcg caa ttt ttg tac ctc	460
His Leu Leu Leu Val Ala Ala Gly Met Glu Ala Gln Phe Leu Tyr Leu	
95 100 105	
tat gcc ttg ata tat ttt cta caa tgc atc aac gca tgt aga att att	508
Tyr Ala Leu Ile Tyr Phe Leu Gln Cys Ile Asn Ala Cys Arg Ile Ile	
110 115 120	
atg aga tgt tgg ctt tgt tgg aag tgc aaa tcc aag aac cca tta ctt	556
Met Arg Cys Trp Leu Cys Trp Lys Cys Lys Ser Lys Asn Pro Leu Leu	
125 130 135 140	
tat gat gcc aac tac ttt gtt tgc tgg cac aca cat aac tat gac tac	604
Tyr Asp Ala Asn Tyr Phe Val Cys Trp His Thr His Asn Tyr Asp Tyr	
145 150 155	
tgt ata cca tat aac agt gtc aca gat aca att gtc gtt act gaa ggt	652
Cys Ile Pro Tyr Asn Ser Val Thr Asp Thr Ile Val Val Thr Glu Gly	
160 165 170	
gac ggc att tca aca cca aaa ctc aaa gaa gac tac caa att ggt ggt	700
Asp Gly Ile Ser Thr Pro Lys Leu Lys Glu Asp Tyr Gln Ile Gly Gly	
175 180 185	
tat tct gag gat agg cac tca ggt gtt aaa gac tat gtc gtt gta cat	748
Tyr Ser Glu Asp Arg His Ser Gly Val Lys Asp Tyr Val Val Val His	
190 195 200	
ggc tat ttc acc gaa gtt tac tac cag ctt gag tct aca caa att act	796
Gly Tyr Phe Thr Glu Val Tyr Tyr Gln Leu Glu Ser Thr Gln Ile Thr	
205 210 215 220	

aca gac act ggt att gaa aat gct aca ttc ttc atc ttt aac aag ctt 844
 Thr Asp Thr Gly Ile Glu Asn Ala Thr Phe Phe Ile Phe Asn Lys Leu
 225 230 235

gtt aaa gac cca ccg aat gtg caa ata cac aca atc gac ggc tct tca 892
 Val Lys Asp Pro Pro Asn Val Gln Ile His Thr Ile Asp Gly Ser Ser
 240 245 250

gga gtt gct aat cca gca atg gat cca att tat gat gag ccg acg acg 940
 Gly Val Ala Asn Pro Ala Met Asp Pro Ile Tyr Asp Glu Pro Thr Thr
 255 260 265

act act agc gtg cct ttg taagcacaag aaagtgagta cgaacttatg 988
 Thr Thr Ser Val Pro Leu
 270

tactcattcg tttcgggaaga aacaggtacg ttaatagttta atagcgtact tcttttttctt 1048

gcttttcgtgg tattcttgct agtcacacta gccatcctta ctgctgctt 1096

<210> 10
 <211> 274
 <212> PRT
 <213> CORONAVIRUS

<400> 10

Met Asp Leu Phe Met Arg Phe Phe Thr Leu Gly Ser Ile Thr Ala Gln
 1 5 10 15

Pro Val Lys Ile Asp Asn Ala Ser Pro Ala Ser Thr Val His Ala Thr
 20 25 30

Ala Thr Ile Pro Leu Gln Ala Ser Leu Pro Phe Gly Trp Leu Val Ile
 35 40 45

Gly Val Ala Phe Leu Ala Val Phe Gln Ser Ala Thr Lys Ile Ile Ala
 50 55 60

Leu Asn Lys Arg Trp Gln Leu Ala Leu Tyr Lys Gly Phe Gln Phe Ile
 65 70 75 80

Cys Asn Leu Leu Leu Leu Phe Val Thr Ile Tyr Ser His Leu Leu Leu
 85 90 95

Val Ala Ala Gly Met Glu Ala Gln Phe Leu Tyr Leu Tyr Ala Leu Ile
 100 105 110

Tyr Phe Leu Gln Cys Ile Asn Ala Cys Arg Ile Ile Met Arg Cys Trp
 115 120 125

Leu Cys Trp Lys Cys Lys Ser Lys Asn Pro Leu Leu Tyr Asp Ala Asn
 130 135 140

ES 2 396 127 T3

Tyr Phe Val Cys Trp His Thr His Asn Tyr Asp Tyr Cys Ile Pro Tyr
145 150 155 160

Asn Ser Val Thr Asp Thr Ile Val Val Thr Glu Gly Asp Gly Ile Ser
165 170 175

Thr Pro Lys Leu Lys Glu Asp Tyr Gln Ile Gly Gly Tyr Ser Glu Asp
180 185 190

Arg His Ser Gly Val Lys Asp Tyr Val Val Val His Gly Tyr Phe Thr
195 200 205

Glu Val Tyr Tyr Gln Leu Glu Ser Thr Gln Ile Thr Thr Asp Thr Gly
210 215 220

Ile Glu Asn Ala Thr Phe Phe Ile Phe Asn Lys Leu Val Lys Asp Pro
225 230 235 240

Pro Asn Val Gln Ile His Thr Ile Asp Gly Ser Ser Gly Val Ala Asn
245 250 255

Pro Ala Met Asp Pro Ile Tyr Asp Glu Pro Thr Thr Thr Thr ser Val
260 265 270

Pro Leu

<210> 11
<211> 1096
<212> ADN
<213> CORONAVIRUS

<220>
<221> CDS
<222> (558)..(1019)
<223>

<400> 11
tcttgctttg ttgcatgact agttgttgca gttgcctcaa ggggtgcatgc tcttggtggtt 60
cttgctgcaa gtttgatgag gatgactctg agccagttct caagggtgtc aaattacatt 120
acacataaac gaacttatgg atttgtttat gagatttttt actcttggat caattactgc 180
acagccagta aaaattgaca atgcttctcc tgcaagtact gttcatgcta cagcaacgat 240
accgctacaa gcctcactcc ctttcggatg gcttggttatt ggcgttgcat ttcttgctgt 300


```

ttttcagagc gctaccaaaa taattgcgct caataaaaga tggcagctag ccctttataa 360
gggcttccag ttcatattgca atttactgct gctattttggt accatctatt cacatctttt 420
gcttgtcgct gcaggatagg aggcgcaatt ttgtacctc tatgccttga tatattttct 480
acaatgcatt aacgcattga gaattattat gagatgttgg ctttgttggga agtgcaaatac 540
caagaaccca ttacttt atg atg cca act act ttg ttt gct ggc aca cac 590
                Met Met Pro Thr Thr Leu Phe Ala Gly Thr His
                1                    5                    10

ata act atg act act gta tac cat ata aca gtg tca cag ata caa ttg 638
Ile Thr Met Thr Thr Val Tyr His Ile Thr Val Ser Gln Ile Gln Leu
                15                    20                    25

tcg tta ctg aag gtg acg gca ttt caa cac caa aac tca aag aag act 686
Ser Leu Leu Lys Val Thr Ala Phe Gln His Gln Asn Ser Lys Lys Thr
                30                    35                    40

acc aaa ttg gtg gtt att ctg agg ata ggc act cag gtg tta aag act 734
Thr Lys Leu Val Val Ile Leu Arg Ile Gly Thr Gln Val Leu Lys Thr
                45                    50                    55

atg tcg ttg tac atg gct att tca ccg aag ttt act acc agc ttg agt 782
Met Ser Leu Tyr Met Ala Ile Ser Pro Lys Phe Thr Thr Ser Leu Ser
60                    65                    70                    75

cta cac aaa tta cta cag aca ctg gta ttg aaa atg cta cat tct tca 830
Leu His Lys Leu Leu Gln Thr Leu Val Leu Lys Met Leu His Ser Ser
80                    85                    90

tct tta aca agc ttg tta aag acc cac cga atg tgc aaa tac aca caa 878
Ser Leu Thr Ser Leu Leu Lys Thr His Arg Met Cys Lys Tyr Thr Gln
95                    100                    105

tcg acg gct ctt cag gag ttg cta atc cag caa tgg atc caa ttt atg 926
Ser Thr Ala Leu Gln Glu Leu Leu Ile Gln Gln Trp Ile Gln Phe Met
110                    115                    120

atg agc cga cga cga cta cta gcg tgc ctt tgt aag cac aag aaa gtg 974
Met Ser Arg Arg Arg Leu Leu Ala Cys Leu Cys Lys His Lys Lys Val
125                    130                    135

agt acg aac tta tgt act cat tcg ttt cgg aag aaa cag gta cgt 1019
Ser Thr Asn Leu Cys Thr His Ser Phe Arg Lys Lys Gln Val Arg
140                    145                    150

taatagttaa tagcgactt ctttttcttg ctttcgtgggt attcttgcta gtcacactag 1079
ccatccttac tgcgctt 1096

```

<210> 12
 <211> 154
 <212> PRT
 <213> CORONAVIRUS

<400> 12

```

Met Met Pro Thr Thr Leu Phe Ala Gly Thr His Ile Thr Met Thr Thr
1                    5                    10                    15

```

ES 2 396 127 T3

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

<210> 13
 <211> 332
 <212> ADN
 <213> CORONAVIRUS

<220>
 <221> CDS
 <222> (36)..(263)
 <223>

<400> 13
 tgcctttgta agcacaagaa agtgagtacg aactt atg tac tca ttc gtt tcg 53
 Met Tyr Ser Phe Val Ser
 1 5
 gaa gaa aca ggt acg tta ata gtt aat agc gta ctt ctt ttt ctt gct 101
 Glu Glu Thr Gly Thr Leu Ile Val Asn Ser Val Leu Leu Phe Leu Ala
 10 15 20
 ttc gtg gta ttc ttg cta gtc aca cta gcc atc ctt act gcg ctt cga 149
 Phe Val Val Phe Leu Leu Val Thr Leu Ala Ile Leu Thr Ala Leu Arg

25	30	35	
ttg tgt gcg tac tgc tgc aat att gtt aac gtg agt tta gta aaa cca			197
Leu Cys Ala Tyr Cys Cys Asn Ile Val Asn Val Ser Leu Val Lys Pro			
40	45	50	
acg gtt tac gtc tac tgc cgt gtt aaa aat ctg aac tct tct gaa gga			245
Thr Val Tyr Val Tyr Ser Arg Val Lys Asn Leu Asn Ser Ser Glu Gly			
55	60	65	70
gtt cct gat ctt ctg gtc taaacgaact aactattatt attattctgt			293
Val Pro Asp Leu Leu Val			
75			
ttggaacttt aacattgctt atcatggcag acaacggta			332

<210> 14
 <211> 76
 <212> PRT
 <213> CORONAVIRUS

<400> 14

Met Tyr Ser Phe Val Ser Glu Glu Thr Gly Thr Leu Ile Val Asn Ser	
1 5 10 15	
Val Leu Leu Phe Leu Ala Phe Val Val Phe Leu Leu Val Thr Leu Ala	
20 25 30	
Ile Leu Thr Ala Leu Arg Leu Cys Ala Tyr Cys Cys Asn Ile Val Asn	
35 40 45	
Val Ser Leu Val Lys Pro Thr Val Tyr Val Tyr Ser Arg Val Lys Asn	
50 55 60	
Leu Asn Ser Ser Glu Gly Val Pro Asp Leu Leu Val	
65 70 75	

<210> 15
 <211> 332
 <212> ADN
 <213> CORONAVIRUS

<400> 15

tgcctttgta agcacaagaa agtgagtacg aacttatgta ctcatctggt tcggaagaaa	60
caggtagctt aatagttaat agcgtacttc ttttcttgc tttcgtggta ttcttgctag	120
tcacactagc catccttact gcgcttcgat tgtgtgcgta ctgctgcaat attgttaacg	180
tgagtttagt aaaaccaacg gtttacgtct actcgcgtgt taaaaatctg aactcttctg	240
aaggagtcc tgatcttctg gtctaaacga actaactatt attattattc tgtttggaac	300
tttaacattg cttatcatgg cagacaacgg ta..	332

<210> 16
 <211> 708
 <212> ADN

<213> CORONAVIRUS

<220>

<221> CDS

<222> (41)..(703)

<223>

<400> 16

tattattatt attctgtttg gaactttaac attgcttatc	atg gca gac aac ggt	55
	Met Ala Asp Asn Gly	
	1 5	
act att acc gtt gag gag ctt aaa caa ctc ctg gaa caa tgg aac cta	103	
Thr Ile Thr Val Glu Glu Leu Lys Gln Leu Leu Glu Gln Trp Asn Leu		
	10 15 20	
gta ata ggt ttc cta ttc cta gcc tgg att atg tta cta caa ttt gcc	151	
Val Ile Gly Phe Leu Phe Leu Ala Trp Ile Met Leu Leu Gln Phe Ala		
	25 30 35	
tat tct aat cgg aac agg ttt ttg tac ata ata aag ctt gtt ttc ctc	199	
Tyr Ser Asn Arg Asn Arg Phe Leu Tyr Ile Ile Lys Leu Val Phe Leu		
	40 45 50	
tgg ctc ttg tgg cca gta aca ctt gct tgt ttt gtg ctt gct gct gtc	247	
Trp Leu Leu Trp Pro Val Thr Leu Ala Cys Phe Val Leu Ala Ala Val		
	55 60 65	
tac aga att aat tgg gtg act ggc ggg att gcg att gca atg gct tgt	295	
Tyr Arg Ile Asn Trp Val Thr Gly Gly Ile Ala Ile Ala Met Ala Cys		
	70 75 80 85	
att gta ggc ttg atg tgg ctt agc tac ttc gtt gct tcc ttc agg ctg	343	
Ile Val Gly Leu Met Trp Leu Ser Tyr Phe Val Ala Ser Phe Arg Leu		
	90 95 100	
ttt gct cgt acc cgc tca atg tgg tca ttc aac cca gaa aca aac att	391	
Phe Ala Arg Thr Arg Ser Met Trp Ser Phe Asn Pro Glu Thr Asn Ile		
	105 110 115	
ctt ctc aat gtg cct ctc cgg ggg aca att gtg acc aga ccg ctc atg	439	
Leu Leu Asn Val Pro Leu Arg Gly Thr Ile Val Thr Arg Pro Leu Met		
	120 125 130	
gaa agt gaa ctt gtc att ggt gct gtg atc att cgt ggt cac ttg cga	487	
Glu Ser Glu Leu Val Ile Gly Ala Val Ile Ile Arg Gly His Leu Arg		
	135 140 145	
atg gcc gga cac tcc cta ggg cgc tgt gac att aag gac ctg cca aaa	535	
Met Ala Gly His Ser Leu Gly Arg Cys Asp Ile Lys Asp Leu Pro Lys		

150		155		160		165	
gag atc act gtg gct aca tca cga acg ctt tct tat tac aaa tta gga							583
Glu Ile Thr Val Ala Thr Ser Arg Thr Leu Ser Tyr Tyr Lys Leu Gly							
		170		175		180	
gcg tcg cag cgt gta ggc act gat tca ggt ttt gct gca tac aac cgc							631
Ala Ser Gln Arg Val Gly Thr Asp Ser Gly Phe Ala Ala Tyr Asn Arg							
		185		190		195	
tac cgt att gga aac tat aaa tta aat aca gac cac gcc ggt agc aac							679
Tyr Arg Ile Gly Asn Tyr Lys Leu Asn Thr Asp His Ala Gly Ser Asn							
		200		205		210	
gac aat att gct ttg cta gta cag taagt							708
Asp Asn Ile Ala Leu Leu Val Gln							
		215		220			

<210> 17
 <211> 221
 <212> PRT
 <213> CORONAVIRUS

<400> 17

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

145		150		155		160									
Lys	Asp	Leu	Pro	Lys	Glu	Ile	Thr	Val	Ala	Thr	Ser	Arg	Thr	Leu	Ser
				165					170					175	
Tyr	Tyr	Lys	Leu	Gly	Ala	Ser	Gln	Arg	Val	Gly	Thr	Asp	Ser	Gly	Phe
			180					185					190		
Ala	Ala	Tyr	Asn	Arg	Tyr	Arg	Ile	Gly	Asn	Tyr	Lys	Leu	Asn	Thr	Asp
		195					200					205			
His	Ala	Gly	Ser	Asn	Asp	Asn	Ile	Ala	Leu	Leu	Val	Gln			
	210					215					220				

<210> 18
 <211> 769
 <212> ADN
 <213> CORONAVIRUS

<400> 18

cctgatcttc	tggtctaaac	gaactaacta	ttattattat	tctgtttgga	actttaacat	60
tgcttatcat	ggcagacaac	ggtactatta	ccgttgagga	gcttaaacia	ctcctggaac	120
aatggaacct	agtaataggt	ttcctattcc	tagcctggat	tatgttacta	caatttgcct	180
attctaatacg	gaacagggtt	ttgtacataa	taaagcttgt	tttcctctgg	ctcttggtggc	240
cagtaacact	tgcttggttt	gtgcttgctg	ctgtctacag	aattaattgg	gtgactggcg	300
ggattgcat	tgcaatggct	tgtattgtag	gcttgatgtg	gcttagctac	ttcgttgctt	360
ccttcaggct	gtttgctcgt	acccgctcaa	tgtggtcatt	caaccagaa	acaaacattc	420
ttctcaatgt	gcctctccgg	gggacaattg	tgaccagacc	gctcatggaa	agtgaacttg	480
tcattggtgc	tgtgatcatt	cgtgggtcact	tgcgaatggc	cggacactcc	ctagggcgct	540
gtgacattaa	ggacctgcca	aaagagatca	ctgtggctac	atcacgaacg	ctttcttatt	600
acaaattagg	agcgtcgcag	cgtgtaggca	ctgattcagg	ttttgctgca	tacaaccgct	660
accgtattgg	aaactataaa	ttaaatacag	accacgccgg	tagcaacgac	aatattgctt	720
tgctagtaca	gtaagtgaca	acagatgttt	catcttggtg	acttccagg		769

<210> 19
 <211> 1231
 <212> ADN
 <213> CORONAVIRUS

<400> 19

taccgtattg gaaactataa attaaataca gaccacgccg gtagcaacga caatattgct	60
ttgctagtag agtaagtgac aacagatggt tcattcttggt gacttccagg ttacaatagc	120
agagatattg attatcatta tgaggacttt caggattgct atttggaatc ttgacgttat	180
aataagttca atagtgagac aattatttaa gcctctaact aagaagaatt attcggagtt	240
agatgatgaa gaacctatgg agttagatta tccataaaac gaacatgaaa attattctct	300
tcctgacatt gattgtattt acatcttgcg agctatatca ctatcaggag tgtgttagag	360
gtacgactgt actactaaaa gaaccttgcc catcaggaac atacgagggc aattcaccat	420
ttcacctctt tgctgacaat aaatttgcac taacttgcac tagcacacac tttgcttttg	480
cttgctga cggtactcga catacctatc agctgcgtgc aagatcagtt tcaccaaacc	540
ttttcatcag acaagaggag gttcaacaag agctctactc gccacttttt ctcatgttg	600
ctgctctagt atttttaata ctttgcttca ccattaagag aaagacagaa tgaatgagct	660
cactttaatt gacttctatt tgtgcttttt agcctttctg ctattccttg ttttaataat	720
gcttattata ttttggtttt cactcgaaat ccaggatcta gaagaacctt gtaccaaagt	780
ctaaacgaac atgaaacttc tcattgtttt gacttgtatt tctctatgca gttgcatatg	840
cactgtagta cagcgctgtg catctaataa acctcatgtg cttgaagatc cttgtaaggt	900
acaacactag gggtaatact tatagcactg cttggctttg tgctctagga aagggtttac	960
cttttcatag atggcacact atgggtcaaa catgcacacc taatgttact atcaactgtc	1020
aagatccagc tgggtggtgcg cttatagcta ggtgttggtta ctttcatgaa ggtcaccaaa	1080
ctgctgcatt tagagacgta cttgttgttt taaataaacg aacaaattaa aatgtctgat	1140
aatggacccc aatcaaacca acgtagtgcc ccccgcatca catttggtgg acccacagat	1200
tcaactgaca ataaccagaa tggaggacgc a	1231

<210> 20
 <211> 1242
 <212> ADN
 <213> CORONAVIRUS

<400> 20

gcatacaacc gctaccgtat tggaaactat aaattaaata cagaccacgc cggtagcaac	60
gacaatattg ctttgctagt acagtaagtg acaacagatg tttcatcttg ttgacttcca	120
ggttacaata gcagagatat tgattatcat tatgaggact ttcaggattg ctatttgga	180
tcttgacgtt ataataagtt caatagtgag acagttattt aagcctctaa ctaagaagaa	240
ttattcggag ttagatgatg aagaacctat ggagtttagat tatccataaa acgaacatga	300
aaattattct cttcctgaca ttgattgtat ttacatcttg cgagctatat cactatcagg	360
agtgtgttag aggtacgact gtactactaa aagaaccttg cccatcagga acatacgagg	420

gcaattcacc atttcaccct cttgctgaca ataaatttgc actaacttgc actagcacac 480
acttttgcttt tgcttgctgt gacgggtactc gacataccta tcagctgcgt gcaagatcag 540
tttcacccaaa actttttcatc agacaagagg aggttcaaca agagctctac tcgccacttt 600
ttctcattgt tgctgctcta gtatTTTTaa tacttttgctt caccattaag agaaagacag 660
aatgaatgag ctcactttta ttgacttcta ttttgctttt ttagcctttc tgctattcct 720
tgTTTTaata atgcttatta tattttgggt ttcactcgaa atccaggatc tagaagaacc 780
ttgtacccaaa gtctaaacga acatgaaact tctcattgtt ttgacttgta tttctctatg 840
cagttgcata tgcactgtag tacagcgctg tgcacttaat aaacctcatg tgcttgaaga 900
tccttgtaag gtacaacact aggggtaata cttatagcac tgcttggctt tgtgctctag 960
gaaaggTTTT accttttcat agatggcaca ctatggttca aacatgcaca cctaattgta 1020
ctatcaactg tcaagatcca gctggtggtg cgcttatagc taggtgttgg taccttcatg 1080
aaggtcacca aactgctgca tttagagacg tacttggtgt tttaaataaa cgaacgaatt 1140
aaaatgtctg ataatggacc ccaatcaaac caacgtagtg cccccgcac tacatttgggt 1200
ggacccacag attcaactga caataaccag aatggaggac gc 1242

<210> 21
<211> 1231
<212> ADN
<213> CORONAVIRUS

<220>
<221> CDS
<222> (86)..(274)
<223>

<400> 21

taccgtattg gaaactataa attaaatata gaccacgccg gtagcaacga caatattgct 60
ttgctagtac agtaagtac aacag atg ttt cat ctt gtt gac ttc cag gtt 112
Met Phe His Leu Val Asp Phe Gln Val
1 5
aca ata gca gag ata ttg att atc att atg agg act ttc agg att gct 160
Thr Ile Ala Glu Ile Leu Ile Ile Ile Met Arg Thr Phe Arg Ile Ala
10 15 20 25
att tgg aat ctt gac gtt ata ata agt tca ata gtg aga caa tta ttt 208
Ile Trp Asn Leu Asp Val Ile Ile Ser Ser Ile Val Arg Gln Leu Phe
30 35 40
aag cct cta act aag aag aat tat tcg gag tta gat gat gaa gaa cct 256
Lys Pro Leu Thr Lys Lys Asn Tyr Ser Glu Leu Asp Asp Glu Glu Pro
45 50 55

atg gag tta gat tat cca taaaacgaac atgaaaatta ttctcttcct 304
Met Glu Leu Asp Tyr Pro
60

gacattgatt gtattttacat cttgcgagct atatcactat caggagtgtg ttagaggtac 364
gactgtacta ctaaaagaac cttgcccatc aggaacatac gagggcaatt caccatttca 424
ccctcttgct gacaataaat ttgcactaac ttgcactagc acacactttg cttttgcttg 484
tgctgacggg actcgacata cctatcagct gcgtgcaaga tcagtttcac caaaactttt 544
catcagacaa gaggagggtc aacaagagct ctactcgcca ctttttctca ttgttgctgc 604
tctagtattt ttaatacttt gcttcacat taagagaaag acagaatgaa tgagctcact 664
ttaattgact tctatttgtg ctttttagcc tttctgctat tccttgtttt aataatgctt 724
attatatattt ggttttcact cgaaatccag gatctagaag aaccttgtag caaagtctaa 784
acgaacatga aacttctcat tgttttgact tgtatttctc tatgcagttg catatgcact 844
gtagtacagc gctgtgcatc taataaacct catgtgcttg aagatccttg taagggtacaa 904
cactaggggt aataacttata gcactgcttg gctttgtgct ctaggaaagg ttttaccttt 964
tcatagatgg cacactatgg ttcaaactg cacacctaata gttactatca actgtcaaga 1024
tccagctggg ggtgcgctta tagctaggtg ttggtacctt catgaagggt accaaactgc 1084
tgcatttaga gacgtacttg ttgttttaaa taaacgaaca aattaaaatg tctgataatg 1144
gaccccaatc aaaccaacgt agtgccccc gcattacatt tggtggaccc acagattcaa 1204
ctgacaataa ccagaatgga ggacgca 1231

<210> 22
<211> 63
<212> PRT
<213> CORONAVIRUS

<400> 22

Met Phe His Leu Val Asp Phe Gln Val Thr Ile Ala Glu Ile Leu Ile
1 5 10 15

Ile Ile Met Arg Thr Phe Arg Ile Ala Ile Trp Asn Leu Asp Val Ile
20 25 30

Ile Ser Ser Ile Val Arg Gln Leu Phe Lys Pro Leu Thr Lys Lys Asn
35 40 45

Tyr Ser Glu Leu Asp Asp Glu Glu Pro Met Glu Leu Asp Tyr Pro
50 55 60

<210> 23
<211> 1231
<212> ADN
<213> CORONAVIRUS

<220>
<221> CDS
<222> (285)..(650)

<223>

<400> 23

taccgtattg gaaactataa attaaataca gaccacgccg gtagcaacga caatattgct	60
ttgctagtac agtaagtac aacagatggt tcatcttggt gacttccagg ttacaatagc	120
agagatattg attatcatta tgaggacttt caggattgct atttggaatc ttgacgttat	180
aataagttca atagttagac aattatttaa gcctctaact aagaagaatt attcggagtt	240
agatgatgaa gaacctatgg agtttagatta tccataaaac gaac atg aaa att att	296
Met Lys Ile Ile	
ctc ttc ctg aca ttg att gta ttt aca tct tgc gag cta tat cac tat	344
Leu Phe Leu Thr Leu Ile Val Phe Thr Ser Cys Glu Leu Tyr His Tyr	
5 10 15 20	
cag gag tgt gtt aga ggt acg act gta cta cta aaa gaa cct tgc cca	392
Gln Glu Cys Val Arg Gly Thr Thr Val Leu Leu Lys Glu Pro Cys Pro	
25 30 35	
tca gga aca tac gag ggc aat tca cca ttt cac cct ctt gct gac aat	440
Ser Gly Thr Tyr Glu Gly Asn Ser Pro Phe His Pro Leu Ala Asp Asn	
40 45 50	
aaa ttt gca cta act tgc act agc aca cac ttt gct ttt gct tgt gct	488
Lys Phe Ala Leu Thr Cys Thr Ser Thr His Phe Ala Phe Ala Cys Ala	
55 60 65	
gac ggt act cga cat acc tat cag ctg cgt gca aga tca gtt tca cca	536
Asp Gly Thr Arg His Thr Tyr Gln Leu Arg Ala Arg Ser Val Ser Pro	
70 75 80	
aaa ctt ttc atc aga caa gag gag gtt caa caa gag ctc tac tcg cca	584
Lys Leu Phe Ile Arg Gln Glu Glu Val Gln Gln Glu Leu Tyr Ser Pro	
85 90 95 100	
ctt ttt ctc att gtt gct gct cta gta ttt tta ata ctt tgc ttc acc	632
Leu Phe Leu Ile Val Ala Ala Leu Val Phe Leu Ile Leu Cys Phe Thr	
105 110 115	
att aag aga aag aca gaa tgaatgagct cactttaatt gacttctatt	680
Ile Lys Arg Lys Thr Glu	
120	
tgtgcttttt agcctttctg ctattccttg ttttaataat gcttattata ttttggtttt	740
cactcgaaat ccaggatcta gaagaacctt gtaccaaagt ctaaacgaac atgaaacttc	800
tcattgtttt gacttgattt tctctatgca gttgcatatg cactgtagta cagcgtgtg	860
catctaataa acctcatgtg cttgaagatc cttgtaagggt acaacactag gggtaatact	920
tatagcactg cttggctttg tgctctagga aagggttttac cttttcatag atggcacact	980
atggttcaaa catgcacacc taatgttact atcaactgtc aagatccagc tggtggtgcg	1040
cttatagcta ggtgttggtg ctttcatgaa ggtcacaaa ctgctgcatt tagagacgta	1100
cttggtgttt taaataaacg aacaaattaa aatgtctgat aatggacccc aatcaaacca	1160
acgtagtgcc ccccgcatca catttggtgg acccacagat tcaactgaca ataaccagaa	1220
tggaggacgc a	1231

ES 2 396 127 T3

<210> 24
 <211> 122
 <212> PRT
 <213> CORONAVIRUS

<400> 24

```

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

Leu Tyr His Tyr Gln Glu Cys Val Arg Gly Thr Thr Val Leu Leu Lys
          20           25           30

Glu Pro Cys Pro Ser Gly Thr Tyr Glu Gly Asn Ser Pro Phe His Pro
          35           40           45

Leu Ala Asp Asn Lys Phe Ala Leu Thr Cys Thr Ser Thr His Phe Ala
 50           55           60

Phe Ala Cys Ala Asp Gly Thr Arg His Thr Tyr Gln Leu Arg Ala Arg
65           70           75           80

Ser Val Ser Pro Lys Leu Phe Ile Arg Gln Glu Glu Val Gln Gln Glu
          85           90           95

Leu Tyr Ser Pro Leu Phe Leu Ile Val Ala Ala Leu Val Phe Leu Ile
          100          105          110

Leu Cys Phe Thr Ile Lys Arg Lys Thr Glu
          115          120
  
```

<210> 25
 <211> 1231
 <212> ADN
 <213> CORONAVIRUS

<220>
 <221> CDS
 <222> (650)..(781)
 <223>

<400> 25

taccgtattg gaaactataa attaaataca gaccacgccg gtagcaacga caatattgct	60
ttgctagtag agtaagtgac aacagatggt tcatcttggt gacttccagg ttacaatagc	120
agagatattg attatcatta tgaggacttt caggattgct atttggaatc ttgacgttat	180
aataagttca atagttagac aattatttaa gcctctaact aagaagaatt attcggagtt	240
agatgatgaa gaacctatgg agttagatta tccataaaac gaacatgaaa attattctct	300
tcctgacatt gattgtatitt acatcttgcg agctatatca ctatcaggag tgtgttagag	360
gtacgactgt actactaaaa gaaccttgcc catcaggaac atacgagggc aattcaccat	420
ttcacccctct tgctgacaat aaatttgcac taacttgcac tagcacacac tttgcttttg	480
cttgtgctga cggtagctga catacctatc agctgcgtgc aagatcagtt tcaccaaacc	540
ttttcatcag acaagaggag gttcaacaag agctctactc gccacttttt ctcatgttg	600
ctgctctagt atttttaata ctttgcttca ccattaagag aaagacaga atg aat gag	658
	Met Asn Glu 1
ctc act tta att gac ttc tat ttg tgc ttt tta gcc ttt ctg cta ttc	706
Leu Thr Leu Ile Asp Phe Tyr Leu Cys Phe Leu Ala Phe Leu Leu Phe	
5 10 15	
ctt gtt tta ata atg ctt att ata ttt tgg ttt tca ctc gaa atc cag	754
Leu Val Leu Ile Met Leu Ile Ile Phe Trp Phe Ser Leu Glu Ile Gln	
20 25 30 35	
gat cta gaa gaa cct tgt acc aaa gtc taaacgaaca tgaaacttct	801
Asp Leu Glu Glu Pro Cys Thr Lys Val	
40	
cattgttttg acttgtatitt ctctatgcag ttgcatatgc actgtagtag agcgtgtgc	861
atctaataaaa cctcatgtgc ttgaagatcc ttgtaaggta caacactagg ggtaatactt	921
atagcactgc ttggctttgt gctctaggaa aggttttacc ttttcataga tggcacacta	981
tggttcaaac atgcacacct aatgttacta tcaactgtca agatccagct ggtggtgcgc	1041
ttatagctag gtgttggtac cttcatgaag gtcaccaaac tgctgcattt agagacgtac	1101
ttgtttgtttt aaataaacga acaaattaaa atgtctgata atggacccca atcaaaccaa	1161
cgtagtgccc cccgcattac atttggtgga cccacagatt caactgacaa taaccagaat	1221
ggaggacgca	1231

<210> 26
 <211> 44
 <212> PRT
 <213> CORONAVIRUS
 <400> 26

ES 2 396 127 T3

Met Asn Glu Leu Thr Leu Ile Asp Phe Tyr Leu Cys Phe Leu Ala Phe
1 5 10 15

Leu Leu Phe Leu Val Leu Ile Met Leu Ile Ile Phe Trp Phe Ser Leu
20 25 30

Glu Ile Gln Asp Leu Glu Glu Pro Cys Thr Lys Val
35 40

<210> 27
<211> 1231
<212> ADN
<213> CORONAVIRUS

<220>
<221> CDS
<222> (791)..(907)
<223>

<400> 27

taccgtattg gaaactataa attaaatata gaccacgccg gtagcaacga caatattgct	60
ttgctagtag agtaagtgac aacagatggt tcatcttggt gacttccagg ttacaatagc	120
agagatattg attatcatta tgaggacttt caggattgct atttggaatc ttgacgttat	180
aataagttca atagttagac aattatttaa gcctctaact aagaagaatt attcggagtt	240
agatgatgaa gaacctatgg agttagatta tccataaaac gaacatgaaa attattctct	300
tcctgacatt gattgtatgt acatcttgcg agctatatca ctatcaggag tgtgttagag	360
gtacgactgt actactaaaa gaaccttgcc catcaggaac atacgagggc aattcaccat	420
ttcacctctt tgctgacaat aaatttgcac taacttgcac tagcacacac tttgcttttg	480
cttggtgctga cggtactcga catacctatc agctgcgtgc aagatcagtt tcaccaaacc	540
ttttcatcag acaagaggag gttcaacaag agctctactc gccacttttt ctcatgttg	600
ctgctctagt atttttaata ctttgcttca ccattaagag aaagacagaa tgaatgagct	660
cactttaatt gacttctatt tgtgcttttt agcctttctg ctattccttg ttttaataat	720
gcttattata ttttggtttt cactcgaaat ccaggatcta gaagaacctt gtaccaaagt	780

```

ctaaacgaac atg aaa ctt ctc att gtt ttg act tgt att tct cta tgc      829
      Met Lys Leu Leu Ile Val Leu Thr Cys Ile Ser Leu Cys
      1          5          10

agt tgc ata tgc act gta gta cag cgc tgt gca tct aat aaa cct cat      877
Ser Cys Ile Cys Thr Val Val Gln Arg Cys Ala Ser Asn Lys Pro His
      15          20          25

gtg ctt gaa gat cct tgt aag gta caa cac taggggtaat acttatagca      927
Val Leu Glu Asp Pro Cys Lys Val Gln His
      30          35

ctgcttggct ttgtgctcta ggaaagggtt taccttttca tagatggcac actatgggttc      987
aaacatgcac acctaattgtt actatcaact gtcaagatcc agctgggtggt gcgcttatag      1047
ctagggtgttg gtaccttcat gaaggtcacc aaactgctgc atttagagac gtacttgttg      1107
ttttaaataa acgaacaaat taaaatgtct gataatggac cccaatcaaa ccaacgtagt      1167
gccccccgca ttacatttgg tggaccacaca gattcaactg acaataacca gaatggagga      1227
cgca                                                                1231

```

<210> 28
 <211> 39
 <212> PRT
 <213> CORONAVIRUS

<400> 28

```

Met Lys Leu Leu Ile Val Leu Thr Cys Ile Ser Leu Cys Ser Cys Ile
      1          5          10          15

Cys Thr Val Val Gln Arg Cys Ala Ser Asn Lys Pro His Val Leu Glu
      20          25          30

Asp Pro Cys Lys Val Gln His
      35

```

<210> 29
 <211> 1231
 <212> ADN
 <213> CORONAVIRUS

<220>
 <221> CDS
 <222> (876)..(1127)
 <223>

<400> 29

```

taccgtattg gaaactataa attaaatata gaccacgccg gtagcaacga caatattgct      60
ttgctagtac agtaagtgac aacagatggt tcattcttgtt gatttccagg ttacaatagc    120
agagatatattg attatcatta tgaggacttt caggattgct atttggaatc ttgacgttat    180
aataagttca atagtgagac aattatttaa gcctctaact aagaagaatt attcggagtt    240
agatgatgaa gaacctatgg agttagatta tccataaaac gaacatgaaa attattctct    300
tcctgacatt gattgtattt acatcttgcg agctatatca ctatcaggag tgtgttagag    360
gtacgactgt actactaaaa gaaccttgcc catcaggaac atacgagggc aattcaccat    420
ttcacctctt tgctgacaat aaatttgcac taacttgcac tagcacacac tttgcttttg    480
cttgtgctga cggctactga catacctatc agctgcgtgc aagatcagtt tcaccaaacc    540
ttttcatcag acaagaggag gttcaacaag agctctactc gccacttttt ctattgttg      600
ctgctctagt atttttaata ctttgcttca ccattaagag aaagacagaa tgaatgagct    660
cactttaatt gacttctatt tgtgcttttt agcctttctg ctattccttg ttttaataat    720
gcttattata ttttggtttt cactcgaaat ccaggatcta gaagaacctt gtaccaaagt    780
ctaaacgaac atgaaacttc tcattgtttt gacttgtatt tctctatgca gttgcatatg    840
cactgtagta cagcgctgtg catctaataa acctc atg tgc ttg aag atc ctt      893
                               Met Cys Leu Lys Ile Leu
                               1                               5

gta agg tac aac act agg ggt aat act tat agc act gct tgg ctt tgt      941
Val Arg Tyr Asn Thr Arg Gly Asn Thr Tyr Ser Thr Ala Trp Leu Cys
                               10                               15                               20

gct cta gga aag gtt tta cct ttt cat aga tgg cac act atg gtt caa      989
Ala Leu Gly Lys Val Leu Pro Phe His Arg Trp His Thr Met Val Gln
                               25                               30                               35

aca tgc aca cct aat gtt act atc aac tgt caa gat cca gct ggt ggt    1037
Thr Cys Thr Pro Asn Val Thr Ile Asn Cys Gln Asp Pro Ala Gly Gly
                               40                               45                               50

gcg ctt ata gct agg tgt tgg tac ctt cat gaa ggt cac caa act gct    1085
Ala Leu Ile Ala Arg Cys Trp Tyr Leu His Glu Gly His Gln Thr Ala
                               55                               60                               65                               70

gca ttt aga gac gta ctt gtt gtt tta aat aaa cga aca aat      1127
Ala Phe Arg Asp Val Leu Val Val Leu Asn Lys Arg Thr Asn
                               75                               80

taaaatgtct gataatggac cccaatcaaa ccaacgtagt gccccccgca ttacatttgg    1187
tggaccacaca gattcaactg acaataacca gaatggagga cgca      1231

```

<210> 30
 <211> 84
 <212> PRT
 <213> CORONAVIRUS

 <400> 30

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

<210> 31
 <211> 21221
 <212> ADN
 <213> CORONAVIRUS

<400> 31
 atggagagcc ttgttcttgg tgtcaacgag aaaacacacg tccaactcag tttgcctgtc 60
 cttcagggtta gagacgtgct agtgcgtggc ttcggggact ctgtggaaga ggccctatcg 120
 gaggcacgtg aacacctcaa aaatggcact tgttgtctag tagagctgga aaaaggcgta 180
 ctgccccagc ttgaacagcc ctatgtgttc attaaacggt ctgatgcctt aagcaccaat 240
 cacggccaca aggtcgttga gctggttgca gaaatggacg gcattcagta cggtcgtagc 300
 ggtataacac tgggagtact cgtgccacat gtgggagaaa cccaattgc ataccgcaat 360
 gttcttcttc gtaagaacgg taataagga gccgggtggc atagctatgg catcgatcta 420
 aagtcttatg acttaggtga cgagcttggc actgatccca ttgaagatta tgaacaaaac 480
 tggaacacta agcatggcag tgggtgcactc cgtgaactca ctcgtgagct caatggaggt 540
 gcagtcactc gctatgtcga caacaatttc tgtggcccag atgggtaccc tcttgattgc 600
 atcaaagatt ttctcgcacg cgcgggcaag tcaatgtgca ctctttccga acaacttgat 660
 tacatcgagt cgaagagagg tgtctactgc tgccgtgacc atgagcatga aattgcctgg 720
 ttactgagc gctctgataa gagctacgag caccagacac ctttcgaaat taagagtgcc 780
 aagaaatttg acactttcaa aggggaatgc ccaaagttag tgtttcctct taactcaaaa 840
 gtcaaagtca ttcaaccacg tgttgaaaag aaaaagactg agggtttcat ggggcgtata 900

cgctctgtgt	accctgttgc	atctccacag	gagtgtaca	atatgcactt	gtctaccttg	960
atgaaatgta	atcattgcga	tgaagtttca	tggcagacgt	gcgactttct	gaaagccact	1020
tgtgaacatt	gtggcactga	aaatttagtt	attgaaggac	ctactacatg	tgggtacctt	1080
cctactaatg	ctgtagtgaa	aatgccatgt	cctgcctgtc	aagaccaga	gattggacct	1140
gagcatagt	ttgcagatta	tcacaaccac	tcaaacattg	aaactcgact	ccgcaaggga	1200
ggtaggacta	gatgttttgg	aggctgtgtg	tttgcctatg	ttggctgcta	taataagcgt	1260
gcctactggg	ttcctcgtgc	tagtgctgat	attggctcag	gccatactgg	cattactggg	1320
gacaatgtgg	agaccttgaa	tgaggatctc	cttgagatac	tgagtcgtga	acgtgttaac	1380
attaacattg	ttggcgattt	tcatttgaat	gaagagggtg	ccatcatttt	ggcatctttc	1440
tctgcttcta	caagtgcctt	tattgacact	ataaagagtc	ttgattacaa	gtctttcaaa	1500
accattgttg	agtcctgcgg	taactataaa	gttaccaagg	gaaagcccgt	aaaagggtgt	1560
tggaaacattg	gacaacagag	atcagtttta	acaccactgt	gtggttttcc	ctcacaggct	1620
gctgggtgtta	tcagatcaat	ttttgcgcgc	acacttgatg	cagcaaacca	ctcaattcct	1680
gatttgcaaa	gagcagctgt	caccatactt	gatgggtattt	ctgaacagtc	attacgtctt	1740
gtcgacgcca	tggtttatac	ttcagacctg	ctcaccaaca	gtgtcattat	tatggcatat	1800
gtaactgggtg	gtcttgtaca	acagacttct	cagtgggtgt	ctaattcttt	gggcactact	1860
gttgaaaaac	tcaggcctat	ctttgaatgg	attgaggcga	aacttagtgc	aggagttgaa	1920
tttctcaagg	atgcttggga	gattctcaaa	tttctcatta	cagggtgttt	tgacatcgtc	1980
aagggtcaaa	tacaggttgc	ttcagataac	atcaaggatt	gtgtaaaatg	cttcattgat	2040
gttgtaaca	aggcactcga	aatgtgcatt	gatcaagtca	ctatcgctgg	cgcaaagttg	2100
cgatcactca	acttaggtga	agtcttcac	gctcaaagca	agggacttta	ccgtcagtgt	2160
atacgtggca	aggagcagct	gcaactactc	atgcctctta	aggcaccaaa	agaagtaacc	2220
tttcttgaag	gtgattcaca	tgacacagta	cttacctctg	aggaggttgt	tctcaagaac	2280
gggtgaactcg	aagcactcga	gacgcccgtt	gatagcttca	caaatggagc	tatcgttggc	2340
acaccagtct	gtgtaaatgg	cctcatgctc	ttagagatta	aggacaaaga	acaatactgc	2400
gcattgtctc	ctggtttact	ggctacaaac	aatgtctttc	gcttaaaagg	gggtgcacca	2460
attaaagggtg	taacctttgg	agaagatact	gtttgggaag	ttcaagggtta	caagaatgtg	2520
agaatcacat	ttgagcttga	tgaacgtgtt	gacaaagtgc	ttaatgaaaa	gtgctctgtc	2580
tacactgttg	aatccggtac	cgaagtact	gagtttgc	gtgtttagc	agaggctgtt	2640
gtgaagactt	tacaaccagt	ttctgatctc	cttaccaaca	tgggtattga	tcttgatgag	2700
tggagtgtag	ctacattcta	cttatttgat	gatgctgggtg	aagaaaactt	ttcatcacgt	2760
atgtattgtt	ccttttacc	tccagatgag	gaagaagagg	acgatgcaga	gtgtgaggaa	2820
gaagaaattg	atgaaacctg	tgaacatgag	tacggtagag	aggatgatta	tcaagggtctc	2880
cctctggaat	ttgggtgcctc	agctgaaaca	gttcgagttg	aggaagaaga	agaggaagac	2940

tggctggatg	atactactga	gcaatcagag	attgagccag	aaccagaacc	tacacctgaa	3000
gaaccagtta	atcagttttac	tggttattta	aaacttactg	acaatggtgc	cattaaatgt	3060
gttgacatcg	ttaaggaggc	acaaagtgct	aatcctatgg	tgattgtaaa	tgctgctaac	3120
atacacctga	aacatggtgg	tggtgtagca	ggtgcactca	acaaggcaac	caatggtgcc	3180
atgcaaaagg	agagtgatga	ttacattaag	ctaaatggcc	ctcttacagt	aggagggtct	3240
tgtttgcttt	ctggacataa	tcttgctaag	aagtgtctgc	atgttggttg	acctaacctt	3300
aatgcagggtg	aggacatcca	gcttcttaag	gcagcatatg	aaaatttcaa	ttcacaggac	3360
atcttacttg	caccattggt	gtcagcaggc	atatttggtg	ctaaaccact	tcagtcttta	3420
caagtgtgcg	tgacagacgg	tcgtacacag	gtttatatattg	cagtcaatga	caaagctctt	3480
tatgagcagg	ttgtcatgga	ttatcttgat	aacctgaagc	ctagagtggg	agcacctaaa	3540
caagaggagc	caccaaacc	agaagattcc	aaaactgagg	agaaatctgt	cgtacagaag	3600
cctgtcgaatg	tgaagccaaa	aattaaggcc	tgcatatgatg	aggttaccac	aacactggaa	3660
gaaactaagt	ttcttaccaa	taagttactc	ttgtttgctg	atatcaatgg	taagctttac	3720
catgattctc	agaacatgct	tagaggtgaa	gatatgtctt	tccttgagaa	ggatgcacct	3780
tacatggtag	gtgatgttat	cactagtggg	gatatcactt	gtgttgtaat	accctccaaa	3840
aaggctgggtg	gcactactga	gatgctctca	agagctttga	agaaagtgcc	agttgatgag	3900
tatataacca	cgtaccctgg	acaaggatgt	gctgggtata	cacttgagga	agctaagact	3960
gctcttaaga	aatgcaaata	tgcatTTTTAT	gtactacctt	cagaagcacc	taatgctaag	4020
gaagagattc	taggaactgt	atcctggaat	ttgagagaaa	tgcttgctca	tgctgaagag	4080
acaagaaaat	taatgcctat	atgcatggat	gttagagcca	taatggcaac	catccaacgt	4140
aagtataaag	gaattaaaat	tcaagagggc	atcgttgact	atggtgtccg	attcttcttt	4200
tatactagta	aagagcctgt	agcttctatt	attacgaagc	tgaactctct	aaatgagccg	4260
cttgtcacia	tgccaattgg	ttatgtgaca	catgggttta	atcttgaaga	ggctgcgcgc	4320
tgtatgcgtt	ctcttaaagc	tcctgccgta	gtgtcagtat	catcaccaga	tgctgttact	4380
acatataatg	gatacctcac	ttcgtcatca	aagacatctg	aggagcactt	tgtagaaaca	4440
gtttcttttg	ctggctctta	cagagattgg	tcctattcag	gacagcgtac	agagttaggt	4500
gttgaatttc	ttaagcgtgg	tgacaaaatt	gtgtaccaca	ctctggagag	ccccgtcgag	4560
tttcatcttg	acggtgaggt	tctttcactt	gacaaactaa	agagtctctt	atccctgcgg	4620
gaggttaaga	ctataaaagt	gttcacaact	gtggacaaca	ctaactctca	cacacagctt	4680
gtggatatgt	ctatgacata	tgacagcag	tttgggtccaa	catacttgga	tggtgctgat	4740
gttacaaaaa	ttaaaccctca	tgtaaatacat	gagggttaaga	ctttctttgt	actacctagt	4800
gatgacacac	tacgtagtga	agctttcgag	tactaccata	ctcttgatga	gagttttctt	4860
ggtaggtaca	tgtctgcttt	aaaccacaca	aagaaatgga	aatttcctca	agttgggtgg	4920
ttaacttcaa	ttaaattgggc	tgataacaat	tgttatttgt	ctagtgtttt	attagcactt	4980

caacagcttg	aagtcaaatt	caatgcacca	gcacttcaag	aggcttatta	tagagcccgt	5040
gctggtgatg	ctgctaactt	ttgtgcactc	atactcgctt	acagtaataa	aactgttggc	5100
gagcttggtg	atgtcagaga	aactatgacc	catcttctac	agcatgctaa	tttggaatct	5160
gcaaagcgag	ttcttaaatgt	ggtgtgtaaa	cattgtggtc	agaaaactac	taccttaacg	5220
ggtgtagaag	ctgtgatgta	tatgggtact	ctatcttatg	ataatcttaa	gacagggtgtt	5280
tccattccat	gtgtgtgtgg	tcgtgatgct	acacaatatc	tagtacaaca	agagtcttct	5340
tttgttatga	tgtctgcacc	acctgctgag	tataaattac	agcaaggtag	attcttatgt	5400
gcgaatgagt	acactggtaa	ctatcagtgt	ggtcattaca	ctcatataac	tgctaaggag	5460
acctctatc	gtattgacgg	agctcacctt	acaaagatgt	cagagtacaa	aggaccagt	5520
actgatgttt	tctacaagga	aacatcttac	actacaacca	tcaagcctgt	gtcgtataaa	5580
ctcgatggag	ttacttacac	agagattgaa	ccaaaattgg	atgggtatta	taaaaaggat	5640
aatgcttact	atacagagca	gcctatagac	cttgtagcaa	ctcaaccatt	accaaattgcg	5700
agttttgata	atttcaaact	cacatgttct	aacacaaaat	ttgctgatga	tttaaataca	5760
atgacaggct	tcacaaagcc	agcttcacga	gagctatctg	tcacattctt	cccagacttg	5820
aatggcgatg	tagtggctat	tgactataga	cactattcag	cgagtttcaa	gaaagggtgct	5880
aaattactgc	ataagccaat	tgtttggcac	attaaccagg	ctacaaccaa	gacaacgttc	5940
aaaccaaaca	cttgggtgttt	acgttgtctt	tggagtacaa	agccagtaga	tacttcaaata	6000
tcatttgaag	ttctggcagt	agaagacaca	caaggaatgg	acaatcttgc	ttgtgaaagt	6060
caacaacca	cctctgaaga	agtagtggaa	aatcctacca	tacagaagga	agtcatagag	6120
tgtgacgtga	aaactaccga	agttgtaggc	aatgtcatac	ttaaaccatc	agatgaaggt	6180
gttaaagtaa	cacaagagtt	aggatcatgag	gatcttatgg	ctgcttatgt	ggaaaacaca	6240
agcattacca	ttaagaaacc	taatgagctt	tcactagcct	taggtttaaa	aacaattgcc	6300
actcatggta	ttgctgcaat	taatagtgtt	ccttggagta	aaattttggc	ttatgtcaaa	6360
ccattcttag	gacaagcagc	aattacaaca	tcaaattgcg	ctaagagatt	agcacaacgt	6420
gtgtttaaca	attatatgcc	ttatgtgttt	acattattgt	tccaattgtg	tacttttact	6480
aaaagtacca	attctagaat	tagagcttca	ctacctacaa	ctattgctaa	aaatagtgtt	6540
aagagtgttg	ctaaattatg	tttggatgcc	ggcattaatt	atgtgaagtc	acccaaattt	6600
tctaaattgt	tcacaatcgc	tatgtggcta	ttgttggtta	gtatttgctt	aggttctcta	6660
atctgtgtaa	ctgctgcttt	tggtgtactc	ttatctaatt	ttgggtgctcc	ttcttattgt	6720
aatggcggtta	gagaattgta	tcttaattcg	tctaacgtta	ctactatgga	tttctgtgaa	6780
ggttcttttc	cttgcagcat	ttgtttaagt	ggattagact	cccttgattc	ttatccagct	6840
cttgaaacca	ttcagggtgac	gatttcatcg	tacaagctag	acttgacaat	tttaggtctg	6900
gccgctgagt	gggttttggc	atatatgttg	ttcacaaaat	tctttttattt	attaggtctt	6960
tcagctataa	tgcagggtgtt	ctttggctat	tttgctagtc	atttcatcag	caattcttgg	7020

ctcatgtggt	ttatcattag	tattgtacaa	atggcaccgc	tttctgcaat	ggttaggatg	7080
tacatcttct	ttgcttcttt	ctactacata	tggaagagct	atgttcatat	catggatggt	7140
tgcacctctt	cgacttgcat	gatgtgctat	aagcgcaatc	gtgccacacg	cgttgagtgt	7200
acaactattg	ttaatggcat	gaagagatct	ttctatgtct	atgcaaattg	aggccgtggc	7260
ttctgcaaga	ctcacaattg	gaattgtctc	aattgtgaca	cattttgcac	tggtagtaca	7320
ttcattagtg	atgaagttgc	tcgtgatttg	tcactccagt	ttaaaagacc	aatcaaccct	7380
actgaccagt	catcgtatat	tgttgatagt	gttgctgtga	aaaatggcgc	gcttcacctc	7440
tactttgaca	aggctgggtca	aaagacctat	gagagacatc	cgctctccca	ttttgtcaat	7500
ttagacaatt	tgagagctaa	caacactaaa	ggttcactgc	ctattaatgt	catagttttt	7560
gatggcaagt	ccaaatgcga	cgagtctgct	tctaagtctg	cttctgtgta	ctacagtcag	7620
ctgatgtgcc	aacctattct	gttgcttgac	caagctcttg	tatcagacgt	tggtgatagt	7680
actgaagttt	ccgttaagat	gtttgatgct	tatgtcgaca	ccttttcagc	aacttttagt	7740
gttcctatgg	aaaaacttaa	ggcacttggt	gctacagctc	acagcgagtt	agcaaagggg	7800
gtagcttttag	atgggtgtcct	ttctacattc	gtgtcagctg	cccgacaagg	tggtgttgat	7860
accgatgttg	acacaaaagga	tgttattgaa	tgtctcaaac	tttcacatca	ctctgactta	7920
gaagtgcag	gtgacagttg	taacaatttc	atgctcacct	ataataagg	tgaaaacatg	7980
acgcccagag	atcttgggcg	atgtattgac	tgtaatgcaa	ggcatatcaa	tgcccaagta	8040
gcaaaaagtc	acaatgtttc	actcatctgg	aatgtaaaag	actacatgtc	tttatctgaa	8100
cagctgcgta	aacaaattcg	tagtgctgcc	aagaagaaca	acataccttt	tagactaact	8160
tgtgctacaa	ctagacaggt	tgtcaatgtc	ataactacta	aaatctcact	caaggggtgg	8220
aagattgtta	gtacttgttt	taaacttatg	cttaaggcca	cattattgtg	cgttcttgct	8280
gcattggttt	gttatatcgt	tatgccagta	catacattgt	caatccatga	tggttacaca	8340
aatgaaatca	ttggttacaa	agccattcag	gatgggtgtca	ctcgtgacat	catttctact	8400
gatgattggt	ttgcaaataa	acatgctggg	tttgacgcat	ggtttagcca	gcgtgggtgg	8460
tcatacaaaa	atgacaaaag	ctgccctgta	gtagctgcta	tcattacaag	agagattggg	8520
ttcatagtc	ctggcttacc	gggtactgtg	ctgagagcaa	tcaatgggtga	cttcttgcat	8580
tttctacctc	gtgttttttag	tgctgttggc	aacatttgct	acacaccttc	caaactcatt	8640
gagtatagtg	attttgctac	ctctgcttgc	gttcttgctg	ctgagtgtac	aatttttaag	8700
gatgctatgg	gcaaacctgt	gccatattgt	tatgacacta	atttgctaga	gggttctatt	8760
tcttatagtg	agcttcgtcc	agacactcgt	tatgtgctta	tggtatgggtc	catcatacag	8820
tttcctaaca	cttacctgga	gggttctggt	agagtagtaa	caacttttga	tgctgagtac	8880
tgtagacatg	gtacatgcga	aaggtcagaa	gtaggtattt	gcctatctac	cagtggtaga	8940
tgggttctta	ataatgagca	ttacagagct	ctatcaggag	ttttctgtgg	tggtgatgcg	9000
atgaatctca	tagctaacat	ctttactcct	cttgtgcaac	ctgtgggtgc	tttagatgtg	9060

tctgcttcag	tagtggctgg	tggtattatt	gccatattgg	tgacttggtc	tgccactac	9120
tttatgaaat	tcagacgtgt	ttttggtgag	tacaaccatg	ttgttgctgc	taatgcactt	9180
ttgtttttga	tgtctttcac	tatactctgt	ctggtaccag	cttacagctt	tctgccggga	9240
gtctactcag	tcttttactt	gtacttgaca	ttctattttca	ccaatgatgt	ttcattcttg	9300
gctcaccttc	aatggtttgc	catgttttct	cctattgtgc	ctttttggat	aacagcaatc	9360
tatgtattct	gtatttctct	gaagcactgc	cattggttct	ttaacaacta	tcttaggaaa	9420
agagtcatgt	ttaatggagt	tacatttagt	accttcgagg	aggctgcttt	gtgtaccttt	9480
ttgctcaaca	aggaaatgta	cctaaaattg	cgtagcgaga	cactgttgcc	acttacacag	9540
tataacaggt	atcttgctct	atataacaag	tacaagtatt	tcagtggagc	cttagatact	9600
accagctatc	gtgaagcagc	ttgctgccac	ttagcaaagg	ctctaaatga	ctttagcaac	9660
tcagggtgctg	atgttctcta	ccaaccacca	cagacatcaa	tcacttctgc	tgttctgcag	9720
agtggtttta	ggaaaatggc	attcccgtca	ggcaaagttg	aaggggtgcat	ggtacaagta	9780
acctgtggaa	ctacaactct	taatggattg	tggttggaatg	acacagtata	ctgtccaaga	9840
catgtcattt	gcacagcaga	agacatgctt	aatcctaact	atgaagatct	gctcattcgc	9900
aaatccaacc	atagctttct	tgttcaggct	ggcaatgttc	aacttcgtgt	tattggccat	9960
tctatgcaaa	attgtctgct	taggcttaaa	gttgatactt	ctaaccctaa	gacaccaag	10020
tataaatttg	tccgtatcca	acctgggtcaa	acattttcag	ttctagcatg	ctacaatggt	10080
tcaccatctg	gtgtttatca	gtgtgccatg	agacctaatc	ataccattaa	aggttctttc	10140
cttaatggat	catgtggtag	tgttggtttt	aacattgatt	atgattgcgt	gtctttctgc	10200
tatatgcata	atatggagct	tccaacagga	gtacacgctg	gtactgactt	agaaggtaaa	10260
ttctatggtc	catttgttga	cagacaaact	gcacaggctg	caggtagaga	cacaaccata	10320
acattaaatg	ttttggcatg	gctgtatgct	gctgttatca	atgggtgatag	gtggtttctt	10380
aatagattca	ccactacttt	gaatgacttt	aaccttggtg	caatgaagta	caactatgaa	10440
cctttgacac	aagatcatgt	tgacatattg	ggacctcttt	ctgctcaaac	aggaattgcc	10500
gtcttagata	tgtgtgctgc	tttgaaagag	ctgctgcaga	atgggtatgaa	tggtcgtact	10560
atccttggtg	gcactatttt	agaagatgag	tttacaccat	ttgatgttgt	tagacaatgc	10620
tctgggtgta	ccttccaagg	taagttcaag	aaaattgtta	agggcactca	tcattggatg	10680
cttttaactt	tcttgacatc	actattgatt	cttggtcaaa	gtacacagtg	gtcactgttt	10740
ttctttgttt	acgagaatgc	tttcttgcca	tttactcttg	gtattatggc	aattgctgca	10800
tgtgctatgc	tgcttgtaaa	gcataagcac	gcattcttgt	gcttgtttct	gttaccttct	10860
cttgcaacag	ttgcttactt	taatatgggtc	tacatgcctg	ctagctgggt	gatgcgtatc	10920
atgacatggc	ttgaattggc	tgacactagc	ttgtctgggt	ataggcttaa	ggattgtgtt	10980
atgtatgctt	cagcttttagt	tttgcttatt	ctcatgacag	ctcgactgt	ttatgatgat	11040
gctgctagac	gtgtttggac	actgatgaat	gtcattacac	ttgtttacaa	agtctactat	11100

ggtaatgctt tagatcaagc tatttccatg tgggccttag ttatttctgt aacctctaac 11160
 tattctggtg tcgttacgac tatcatgttt ttagctagag ctatagtgtt tgtgtgtgtt 11220
 gagtattacc cattgttatt tattactggc aacaccttac agtgtatcat gcttgtttat 11280
 tgtttcttag gctattgttg ctgctgctac tttggccttt tctgtttact caaccgttac 11340
 ttcaggctta ctcttggtgt ttatgactac ttggtctcta cacaagaatt taggtatatg 11400
 aactccagg ggcttttgcc tcctaagagt agtattgatg ctttcaagct taacattaag 11460
 ttgttgggta ttggaggtaa accatgtatc aagggttgcta ctgtacagtc taaaatgtct 11520
 gacgtaaagt gcacatctgt ggtactgctc tcggttcttc aacaacttag agtagagtca 11580
 tcttctaaat tgtgggcaca atgtgtacaa ctccacaatg atattcttct tgcaaaagac 11640
 acaactgaag ctttcgagaa gatggtttct cttttgtctg ttttgctatc catgcagggt 11700
 gctgtagaca ttaataggtt gtgcgaggaa atgctcgata accgtgctac tcttcaggct 11760
 attgcttcag aatttagttc ttaccatca tatgccgctt atgccactgc ccaggaggcc 11820
 tatgagcagg ctgtagctaa tgggtattct gaagtcgttc tcaaaaagtt aaagaaatct 11880
 ttgaatgtgg ctaaactctga gtttgaccgt gatgctgcca tgcaacgcaa gttggaaaag 11940
 atggcagatc aggctatgac ccaaattgtac aaacaggcaa gatctgagga caagagggca 12000
 aaagtaacta gtgctatgca aacaatgctc ttcactatgc ttaggaagct tgataatgat 12060
 gcacttaaca acattatcaa caatgcgcgt gatggttgtg ttccactcaa catcatacca 12120
 ttgactacag cagccaaact catggttgtt gtccctgatt atggtacct caagaacact 12180
 tgtgatggta acacctttac atatgcatct gcactctggg aaatccagca agttgttgat 12240
 gcggatagca agattgttca acttagtgaa attaacatgg acaattcacc aaatttggct 12300
 tggcctctta ttgttacagc tctaagagcc aactcagctg ttaaaactaca gaataatgaa 12360
 ctgagtccag tagcactacg acagatgtcc tgtgcggctg gtaccacaca aacagcttgt 12420
 actgatgaca atgcacttgc ctactataac aattcgaagg gaggtagggt tgtgctggca 12480
 ttactatcag accaccaaga tctcaaattg gctagattcc ctaagagtga tggtagaggt 12540
 acaatttaca cagaactgga accaccttgt aggtttgtta cagacacacc aaaagggcct 12600
 aaagtgaat acttgtaact catcaaaggc ttaaacaacc taaatagagg tatggtgctg 12660
 ggcagttag ctgctacagt acgtcttcag gctggaaatg ctacagaagt acctgccaat 12720
 tcaactgtgc tttccttctg tgcttttgca gtagaccctg ctaaagcata taaggattac 12780
 ctagcaagtg gaggacaacc aatcaccaac tgtgtgaaga tgttgtgtac acacactggt 12840
 acaggacagg caattactgt aacaccagaa gctaacatgg accaagagtc ctttgggtgt 12900
 gcttcatgtt gtctgtattg tagatgccac attgaccatc caaatcctaa aggattctgt 12960
 gacttgaaag gtaagtacgt ccaaatacct accacttgtg ctaatgacc agtgggtttt 13020
 acacttagaa acacagtctg taccgtctgc ggaatgtgga aagggttatg ctgtagtgt 13080
 gaccaactcc gcgaaccctt gatgcagtct gcggatgcat caacgttttt aaacgggttt 13140

gcggtgtaag tgcagcccgt cttacaccgt gcggcacagg cactagtact gatgtcgtct 13200
 acagggcctt tgatatttac aacgaaaaag ttgctgggtt tgcaaagttc ctaaaaacta 13260
 attgctgtcg cttccaggag aaggatgagg aaggcaattt attagactct tactttgtag 13320
 ttaagaggca tactatgtct aactaccaac atgaagagac tatttataac ttgggttaaag 13380
 attgtccagc ggttgctgtc catgactttt tcaagtttag agtagatggt gacatgggtac 13440
 cacatatatc acgtcagcgt ctaactaaat acacaatggc tgatttagtc tatgctctac 13500
 gtcattttga tgagggtaat tgtgatacat taaaagaaat actcgtcaca tacaattgct 13560
 gtgatgatga ttatttcaat aagaaggatt ggtatgactt cgtagagaat cctgacatct 13620
 tacgcgata tgctaactta ggtgagcgtg tacgccaatc attattaaag actgtacaat 13680
 tctgcgatgc tatgcgtgat gcaggcattg taggcgtact gacattagat aatcaggatc 13740
 ttaatgggaa ctggtacgat ttcggtgatt tcgtacaagt agcaccaggc tgcggagttc 13800
 ctattgtgga ttcataattac tcattgctga tgcccatcct cactttgact agggcattgg 13860
 ctgctgagtc ccatatggat gctgatctcg caaaaccact tattaagtgg gatttgctga 13920
 aatatgattt tacggaagag agactttgtc tcttcgaccg ttattttaaa tattgggacc 13980
 agacatacca tccaattgt attaactggt tggatgatag gtgtatcctt cattgtgcaa 14040
 actttaatgt gttattttct actgtgtttc cacctacaag ttttggacca ctagtaagaa 14100
 aaatatttgt agatggtgtt cttttgttg tttcaactgg ataccatttt cgtgagttag 14160
 gagtcgtaca taatcaggat gtaaacttac atagctcgcg tctcagtttc aaggaacttt 14220
 tagtgtatgc tgctgatcca gctatgcatg cagcttctgg caattttattg ctagataaac 14280
 gcactacatg cttttcagta gctgcactaa caaacaatgt tgcttttcaa actgtcaaac 14340
 ccggtaatth taataaagac ttttatgact ttgctgtgtc taaaggtttc tttaaggaag 14400
 gaagtctgtg tgaactaaaa cacttcttct ttgctcagga tggcaacgct gctatcagtg 14460
 attatgacta ttatcgttat aatctgccaa caatgtgtga tatcagacaa ctctatttcg 14520
 tagttgaagt tggtgataaa tacttttgatt gttacgatgg tggctgtatt aatgccaacc 14580
 aagtaatcgt taacaatctg gataaatcag ctggtttccc atttaataaa tggggtaagg 14640
 ctagacttta ttatgactca atgagttatg aggatcaaga tgcacttttc gcgtatacta 14700
 agcgtaatgt catccctact ataactcaaa tgaatcttaa gtatgccatt agtgcaaaga 14760
 atagagctcg caccgtagct ggtgtctcta tctgtagtac tatgacaaat agacagtttc 14820
 atcagaaatt attgaagtca atagccgcca ctagaggagc tactgtggta attggaacaa 14880
 gcaagtttta cggtggctgg cataatatgt taaaaactgt ttacagtgat gtagaaactc 14940
 cacaccttat gggttgggat tatccaaaat gtgacagagc catgcctaac atgcttagga 15000
 taatggcctc tcttgttctt gctcgcaaac ataacacttg ctgtaactta tcacaccgtt 15060
 tctacagggt agctaacgag tgtgcgcaag tattaagtga gatggcatg tgtggcggct 15120
 cactatatgt taaaccagggt ggaacatcat ccggtgatgc tacaactgct tatgctaata 15180

gtgtcttttaa catttgtcaa gctgtttacag ccaatgtaaa tgcacttctt tcaactgatg 15240
 gtaataagat agctgacaag tatgtccgca atctacaaca caggctctat gagtgtctct 15300
 atagaaatag ggatgttgat catgaattcg tggatgagtt ttacgcttac ctgcgtaaac 15360
 atttctccat gatgattctt tctgatgatg ccgttgtgtg ctataacagt aactatgcgg 15420
 ctcaagggtt agtagctagc attaagaact ttaaggcagt tctttattat caaaataatg 15480
 tgttcatgtc tgaggcaaaa tgttggactg agactgacct tactaaagga cctcacgaat 15540
 ttgtctcaca gcatacaatg ctagttaaac aaggagatga ttacgtgtac ctgccttacc 15600
 cagatccatc aagaatatta ggcgcaggct gttttgtcga tgatattgtc aaaacagatg 15660
 gtacacttat gattgaaagg ttcgtgtcac tggctattga tgcttacca cttacaaaac 15720
 atcctaataca ggagtatgct gatgtctttc acttgtattt acaatacatt agaaagttac 15780
 atgatgagct tactggccac atgttggaca tgtattccgt aatgctaact aatgataaca 15840
 cctcacggta ctgggaacct gagttttatg aggctatgta cacaccacat acagtcttgc 15900
 aggctgtagg tgcttgtgta ttgtgcaatt cacagacttc acttcgttgc ggtgcctgta 15960
 ttaggagacc attcctatgt tgcaagtgtc gctatgacca tgtcatttca acatcacaca 16020
 aattagtgtt gtctgttaat ccctatgttt gcaatgcccc aggttgtgat gtcactgatg 16080
 tgacacaact gtatctagga ggtatgagct attattgcaa gtcacataag cctcccatta 16140
 gttttccatt atgtgctaata ggtcagggtt ttggtttata caaaaacaca tgtgtaggca 16200
 gtgacaatgt cactgacttc aatgcgatag caacatgtga ttggactaat gctggcgatt 16260
 acataacttgc caacacttgt actgagagac tcaagctttt cgcagcagaa acgctcaaag 16320
 ccactgagga aacaatttaag ctgtcatatg gtattgccac tgtacgcgaa gtactctctg 16380
 acagagaatt gcatctttca tgggagggtg gaaaacctag accaccattg aacagaaact 16440
 atgtctttac tgggtaccgt gtaactaaaa atagtaaagt acagattgga gagtacacct 16500
 ttgaaaaagg tgactatggt gatgctgttg tgtacagagg tactacgaca tacaagttga 16560
 atgttgggtga ttactttgtg ttgacatctc aactgtaat gccacttagt gcacctactc 16620
 tagtgccaca agagcactat gtgagaatta ctggcttgta cccaacactc aacatctcag 16680
 atgagttttc tagcaatgtt gcaaattatc aaaaggctcg catgcaaaag tactctacac 16740
 tccaaggacc acctggtact ggtaagagtc attttgccat cggacttgct ctctattacc 16800
 catctgctcg catagtgtat acggcatgct ctcatgcagc tgttgatgcc ctatgtgaaa 16860
 aggcattaaa atat ttgccc atagataaat gtagtagaat catacctgcg cgtgcgcgcg 16920
 tagagtgttt tgataaattc aaagtgaatt caacactaga acagtatgtt ttctgcactg 16980
 taaatgcatt gccagaaaca actgctgaca ttgtagtctt tgatgaaatc tctatggcta 17040
 ctaattatga cttgagtgtt gtcaatgcta gacttcgtgc aaaacactac gtctatattg 17100
 gcgatcctgc tcaattacca gcccccgca cattgctgac taaaggcaca ctagaaccag 17160
 aatatttttaa ttcagtgtgc agacttatga aaacaatagg tccagacatg ttccttggaa 17220

cttgtcgccg ttgtcctgct gaaattgttg acactgtgag tgcttttagtt tatgacaata 17280
 agctaaaagc acacaaggat aagtcagctc aatgcttcaa aatgttctac aaagggtgta 17340
 ttacacatga tgtttcatct gcaatcaaca gacctcaaat aggcgttgta agagaatttc 17400
 ttacacgcaa tcctgcttgg agaaaagctg tttttatctc accttataat tcacagaacg 17460
 ctgtagcttc aaaaatctta ggattgccta cgcagactgt tgattcatca cagggttctg 17520
 aatatgacta tgtcatattc acacaaacta ctgaaacagc acactcttgt aatgtcaacc 17580
 gcttcaatgt ggctatcaca agggcaaaaa ttggcatttt gtgcataatg tctgatagag 17640
 atctttatga caaactgcaa tttaacaagtc tagaaatacc acgtcgcaat gtggctacat 17700
 tacaagcaga aaatgtaact ggacttttta aggactgtag taagatcatt actggtcttc 17760
 atcctacaca ggcacctaca cacctcagcg ttgatataaa gttcaagact gaaggattat 17820
 gtgttgacat accaggcata ccaaaggaca tgacctaccg tagactcatc tctatgatgg 17880
 gtttcaaaat gaattaccaa gtcaatgggt accctaatat gtttatcacc cgcgaagaag 17940
 ctattcgtca cgttcgtgcg tggattggct ttgatgtaga gggctgtcat gcaactagag 18000
 atgctgtggg tactaaccta cctctccagc taggattttc tacagggtgtt aacttagtag 18060
 ctgtaccgac tggttatgtt gacactgaaa ataacacaga attcaccaga gttaatgcaa 18120
 aacctccacc aggtgaccag tttaaacatc ttataccact catgtataaa ggcttgccct 18180
 ggaatgtagt gcgtattaag atagtacaaa tgctcagtga tacactgaaa ggattgtcag 18240
 acagagtcgt gttcgtcctt tgggcgcagt gctttgagct tacatcaatg aagtactttg 18300
 tcaagattgg acctgaaaga acgtgttgct tgtgtgacaa acgtgcaact tgcttttcta 18360
 cttcatcaga tacttatgcc tgctggaatc attctgtggg ttttgactat gtctataacc 18420
 catttatgat tgatgttcag cagtggggct ttacgggtaa ccttcagagt aaccatgacc 18480
 aacattgcca ggtacatgga aatgcacatg tggctagtgt tgatgctatc atgactagat 18540
 gtttagcagt ccatgagtgc tttgttaagc gcgttgattg gtctgttgaa taccctatta 18600
 taggagatga actgaggggt aattctgctt gcagaaaagt acaacacatg gttgtgaagt 18660
 ctgcattgct tgctgataag ttccagttc ttcatgacat tggaaatcca aaggctatca 18720
 agtgtgtgcc tcaggctgaa gtagaatgga agttctacga tgctcagcca tgtagtgaca 18780
 aagcttaciaa aatagaggaa ctcttctatt cttatgctac acatcacgat aaattcactg 18840
 atgggtgttg tttgttttg aattgtaacg ttgatcgta cccagccaat gcaattgtgt 18900
 gtaggtttga cacaagagtc ttgtcaaact tgaacttacc aggctgtgat ggtggtagtt 18960
 tgtatgtgaa taagcatgca ttccacactc cagctttcga taaaagtgca ttactaatt 19020
 taaagcaatt gcctttcttt tactattctg atagtccttg tgagtctcat ggcaaacaag 19080
 tagtgtcgga tattgattat gttccactca aatctgctac gtgtattaca cgatgcaatt 19140
 taggtgggtgc tgtttgcaga caccatgcaa atgagtaccg acagtacttg gatgcatata 19200
 atatgatgat ttctgctgga tttagcctat ggatttacia acaatttgat acttataacc 19260

tgtggaatac	atttaccagg	ttacagagtt	tagaaaatgt	ggcttataat	gttggttaata	19320
aaggacactt	tgatggacac	gccggcgaag	cacctgtttc	catcattaat	aatgctgttt	19380
acacaaaggt	agatgggtatt	gatgtggaga	tctttgaaaa	taagacaaca	cttcctgtta	19440
atgttgcatt	tgagcttttg	gctaagcgta	acattaaacc	agtgccagag	attaagatac	19500
tcaataatth	gggtgttgat	atcgctgcta	atactgtaat	ctgggactac	aaaagagaag	19560
ccccagcaca	tgtatctaca	ataggtgtct	gcacaatgac	tgacattgcc	aagaaaccta	19620
ctgagagtgc	ttgttcttca	cttactgtct	tgtttgatgg	tagagtggaa	ggacaggtag	19680
acctttttag	aaacgcccgt	aatgggtgtt	taataacaga	aggttcagtc	aaaggctctaa	19740
caccttcaaa	gggaccagca	caagctagcg	tcaatggagt	cacattaatt	ggagaatcag	19800
taaaaacaca	gtttaactac	tttaagaaag	tagacggcat	tattcaacag	ttgcctgaaa	19860
cctactttac	tcagagcaga	gacttagagg	attttaagcc	cagatcacia	atggaaaactg	19920
actttctcga	gctcgctatg	gatgaattca	tacagcgata	taagctcgag	ggctatgcct	19980
tcgaacacat	cgtttatgga	gatttcagtc	atggacaact	tggcggctct	catttaatga	20040
taggccttagc	caagcgctca	caagattcac	cacttaaatt	agaggatttt	atccctatgg	20100
acagcacagt	gaaaaattac	ttcataacag	atgcgcaaac	aggttcatca	aaatgtgtgt	20160
gttctgtgat	tgatctttta	cttgatgact	ttgtcgagat	aataaagtca	caagatttgt	20220
cagtgatthc	aaaagtggtc	aaggttacaa	ttgactatgc	tgaaatttca	ttcatgcttt	20280
ggtgtaagga	tggacatggt	gaaaccttct	acccaaaact	acaagcaagt	caagcgtggc	20340
aaccaggtgt	tgcatgctct	aacttgtaca	agatgcaaag	aatgcttctt	gaaaagtgtg	20400
accttcagaa	ttatgggtgaa	aatgctgtta	taccaaagg	aataatgatg	aatgtcgcaa	20460
agtatactca	actgtgtcaa	tacttaaata	cacttacttt	agctgtaccc	tacaacatga	20520
gagttattca	ctttgggtgct	ggctctgata	aaggagttgc	accaggtaca	gctgtgctca	20580
gacaatgggt	gccaaactggc	acactacttg	tcgattcaga	tcttaatgac	ttcgtctccg	20640
acgcagattc	tactttaatt	ggagactgtg	caacagtaca	tacggctaatt	aaatgggacc	20700
ttattattag	cgatatgtat	gaccctagga	caaacaatgt	gacaaaagag	aatgactcta	20760
aagaaggggt	tttcacttat	ctgtgtggat	ttataaagca	aaaactagcc	ctgggtgggt	20820
ctatagctgt	aaagataaca	gagcattctt	ggaatgctga	cctttacaag	cttatgggccc	20880
atttctcatg	gtggacagct	tttgttacaa	atgtaaattgc	atcatcatcg	gaagcatttt	20940
taattggggc	taactatctt	ggcaagccga	aggaacaaat	tgatggctat	accatgcatg	21000
ctaactacat	tttctggagg	aacacaaatc	ctatccagtt	gtcttcctat	tcactctttg	21060
acatgagcaa	atttcctctt	aaattaagag	gaactgctgt	aatgtctctt	aaggagaatc	21120
aatcaatga	tatgatttat	tctcttctgg	aaaaaggtag	gcttatcatt	agagaaaaca	21180
acagagttgt	ggtttcaagt	gatattcttg	ttaacaacta	a		21221

<211> 297
<212> ADN
<213> CORONAVIRUS

<400> 32

atggaccca atcaaacc	cgtagtgccc	cccgattac	atttggtgga	cccacagatt	60
caactgaca	taaccagaat	ggaggacgca	atggggcaag	gccaaaacag	cgccgacccc 120
aaggtttacc	caataatact	gcgtcttggt	tcacagctct	cactcagcat	ggcaaggagg 180
aacttagatt	ccctcgaggc	cagggcggtc	caatcaacac	caatagtggg	ccagatgacc 240
aaattggcta	ctaccgaaga	gctacccgac	gagttcgtgg	tggtgacggc	aaaatga 297

<210> 33
<211> 98
<212> PRT
<213> CORONAVIRUS

<400> 33

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

<210> 34
<211> 213
<212> ADN
<213> CORONAVIRUS

<400> 34

atgctgccac	cgtgctacaa	cttcctcaag	gaacaacatt	gccaaaaggc	ttctacgcag	60
aggggaagcag	aggcggcagt	caagcctctt	ctcgctcctc	atcacgtagt	cgcggttaatt	120
caagaaattc	aactcctggc	agcagtaggg	gaaattctcc	tgctcgaatg	gctagcggag	180
gtgggtgaaac	tgccctcgcg	ctattgctgc	tag			213

<210> 35
<211> 70
<212> PRT

<213> CORONAVIRUS

<400> 35

```

Met Leu Pro Pro Cys Tyr Asn Phe Leu Lys Glu Gln His Cys Gln Lys
 1          5          10          15

Ala Ser Thr Gln Arg Glu Ala Glu Ala Val Lys Pro Leu Leu Ala
          20          25          30

Pro His His Val Val Ala Val Ile Gln Glu Ile Gln Leu Leu Ala Ala
          35          40          45

Val Gly Glu Ile Leu Leu Leu Glu Trp Leu Ala Glu Val Val Lys Leu
          50          55          60

Pro Ser Arg Tyr Cys Cys
65          70

```

<210> 36

<211> 1377

<212> ADN

<213> CORONAVIRUS

<220>

<221> CDS

<222> (67)..(1335)

<223>

<400> 36

```

atgaagggtca ccaaactgct gcatttagag acgtacttgt tgttttaaat aaacgaacaa      60

```

attaa	atg	tct	gat	aat	gga	ccc	caa	tca	aac	caa	cgt	agt	gcc	ccc		108
	Met	Ser	Asp	Asn	Gly	Pro	Gln	Ser	Asn	Gln	Arg	Ser	Ala	Pro		
	1				5					10						
cgc	att	aca	ttt	ggt	gga	ccc	aca	gat	tca	act	gac	aat	aac	cag	aat	156
Arg	Ile	Thr	Phe	Gly	Gly	Pro	Thr	Asp	Ser	Thr	Asp	Asn	Asn	Gln	Asn	
15				20						25					30	
gga	gga	cgc	aat	ggg	gca	agg	cca	aaa	cag	cgc	cga	ccc	caa	ggt	tta	204
Gly	Gly	Arg	Asn	Gly	Ala	Arg	Pro	Lys	Gln	Arg	Arg	Pro	Gln	Gly	Leu	
				35					40					45		
ccc	aat	aat	act	gcg	tct	tgg	ttc	aca	gct	ctc	act	cag	cat	ggc	aag	252
Pro	Asn	Asn	Thr	Ala	Ser	Trp	Phe	Thr	Ala	Leu	Thr	Gln	His	Gly	Lys	
			50					55					60			
gag	gaa	ctt	aga	ttc	cct	cga	ggc	cag	ggc	gtt	cca	atc	aac	acc	aat	300
Glu	Glu	Leu	Arg	Phe	Pro	Arg	Gly	Gln	Gly	Val	Pro	Ile	Asn	Thr	Asn	
		65					70					75				
agt	ggt	cca	gat	gac	caa	att	ggc	tac	tac	cga	aga	gct	acc	cga	cga	348
Ser	Gly	Pro	Asp	Asp	Gln	Ile	Gly	Tyr	Tyr	Arg	Arg	Ala	Thr	Arg	Arg	
	80					85					90					
gtt	cgt	ggt	ggt	gac	ggc	aaa	atg	aaa	gag	ctc	agc	ccc	aga	tgg	tac	396
Val	Arg	Gly	Gly	Asp	Gly	Lys	Met	Lys	Glu	Leu	Ser	Pro	Arg	Trp	Tyr	
95					100					105					110	
ttc	tat	tac	cta	gga	act	ggc	cca	gaa	gct	tca	ctt	ccc	tac	ggc	gct	444
Phe	Tyr	Tyr	Leu	Gly	Thr	Gly	Pro	Glu	Ala	Ser	Leu	Pro	Tyr	Gly	Ala	
				115					120					125		
aac	aaa	gaa	ggc	atc	gta	tgg	gtt	gca	act	gag	gga	gcc	ttg	aat	aca	492
Asn	Lys	Glu	Gly	Ile	Val	Trp	Val	Ala	Thr	Glu	Gly	Ala	Leu	Asn	Thr	
			130					135					140			
ccc	aaa	gac	cac	att	ggc	acc	cgc	aat	cct	aat	aac	aat	gct	gcc	acc	540
Pro	Lys	Asp	His	Ile	Gly	Thr	Arg	Asn	Pro	Asn	Asn	Asn	Ala	Ala	Thr	
		145					150					155				
gtg	cta	caa	ctt	cct	caa	gga	aca	aca	ttg	cca	aaa	ggc	ttc	tac	gca	588
Val	Leu	Gln	Leu	Pro	Gln	Gly	Thr	Thr	Leu	Pro	Lys	Gly	Phe	Tyr	Ala	
	160					165					170					
gag	gga	agc	aga	ggc	ggc	agt	caa	gcc	tct	tct	cgc	tcc	tca	tca	cgt	636
Glu	Gly	Ser	Arg	Gly	Gly	Ser	Gln	Ala	Ser	Ser	Arg	Ser	Ser	Ser	Arg	
175				180						185					190	
agt	cgc	ggt	aat	tca	aga	aat	tca	act	cct	ggc	agc	agt	agg	gga	aat	684
Ser	Arg	Gly	Asn	Ser	Arg	Asn	Ser	Thr	Pro	Gly	Ser	Ser	Arg	Gly	Asn	
				195					200					205		
tct	cct	gct	cga	atg	gct	agc	gga	ggt	ggt	gaa	act	gcc	ctc	gcg	cta	732
Ser	Pro	Ala	Arg	Met	Ala	Ser	Gly	Gly	Gly	Glu	Thr	Ala	Leu	Ala	Leu	
			210					215					220			
ttg	ctg	cta	gac	aga	ttg	aac	cag	ctt	gag	agc	aaa	gtt	tct	ggt	aaa	780
Leu	Leu	Leu	Asp	Arg	Leu	Asn	Gln	Leu	Glu	Ser	Lys	Val	Ser	Gly	Lys	
		225					230					235				
ggc	caa	caa	caa	caa	ggc	caa	act	gtc	act	aag	aaa	tct	gct	gct	gag	828
Gly	Gln	Gln	Gln	Gln	Gly	Gln	Thr	Val	Thr	Lys	Lys	Ser	Ala	Ala	Glu	
	240					245					250					
gca	tct	aaa	aag	cct	cgc	caa	aaa	cgt	act	gcc	aca	aaa	cag	tac	aac	876
Ala	Ser	Lys	Lys	Pro	Arg	Gln	Lys	Arg	Thr	Ala	Thr	Lys	Gln	Tyr	Asn	
255					260					265					270	

gtc act caa gca ttt ggg aga cgt ggt cca gaa caa acc caa gga aat	924
Val Thr Gln Ala Phe 275 Gly Arg Arg Gly Pro 280 Glu Gln Thr Gln Gly Asn 285	
ttc ggg gac caa gac cta atc aga caa gga act gat tac aaa cat tgg	972
Phe Gly Asp Gln Asp Leu Ile Arg Gln 295 Gly Thr Asp Tyr Lys 300 His Trp	
ccg caa att gca caa ttt gct cca agt gcc tct gca ttc ttt gga atg	1020
Pro Gln Ile Ala Gln Phe Ala Pro Ser Ala Ser Ala Phe 315 Phe Gly Met	
tca cgc att ggc atg gaa gtc aca cct tcg gga aca tgg ctg act tat	1068
Ser Arg Ile Gly Met Glu Val 325 Thr Pro Ser Gly Thr 330 Trp Leu Thr Tyr	
cat gga gcc att aaa ttg gat gac aaa gat cca caa ttc aaa gac aac	1116
His Gly Ala Ile Lys Leu 340 Asp Asp Lys Asp Pro 345 Gln Phe Lys Asp Asn 350	
gtc ata ctg ctg aac aag cac att gac gca tac aaa aca ttc cca cca	1164
Val Ile Leu Leu Asn 355 Lys His Ile Asp Ala Tyr Lys Thr Phe Pro 365 Pro	
aca gag cct aaa aag gac aaa aag aaa aag act gat gaa gct cag cct	1212
Thr Glu Pro Lys 370 Lys Asp Lys Lys Lys Lys Thr Asp Glu 380 Ala Gln Pro	
ttg ccg cag aga caa aag aag cag ccc act gtg act ctt ctt cct gcg	1260
Leu Pro Gln Arg Gln Lys Lys Gln 390 Pro Thr Val Thr Leu 395 Leu Pro Ala	
gct gac atg gat gat ttc tcc aga caa ctt caa aat tcc atg agt gga	1308
Ala Asp Met Asp Asp Phe 405 Ser Arg Gln Leu Gln Asn 410 Ser Met Ser Gly	
gct tct gct gat tca act cag gca taa acactcatga tgaccacaca	1355
Ala Ser Ala Asp Ser Thr 420 Gln Ala	
aggcagatgg gctatgtaaa cg	1377

<210> 37

<211> 422

<212> PRT

<213> CORONAVIRUS

<400> 37

Met Ser Asp Asn Gly Pro Gln Ser Asn Gln Arg Ser Ala Pro Arg Ile	
1 5 10 15	
Thr Phe Gly Gly Pro Thr Asp Ser Thr Asp Asn Asn Gln Asn Gly Gly	
20 25 30	
Arg Asn Gly Ala Arg Pro Lys Gln Arg Arg Pro Gln Gly Leu Pro Asn	
35 40 45	
Asn Thr Ala Ser Trp Phe Thr Ala Leu Thr Gln His Gly Lys Glu Glu	

ES 2 396 127 T3

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

325 330 335

Ala Ile Lys Leu Asp Asp Lys Asp Pro Gln Phe Lys Asp Asn Val Ile
340 345 350

Leu Leu Asn Lys His Ile Asp Ala Tyr Lys Thr Phe Pro Pro Thr Glu
355 360 365

Pro Lys Lys Asp Lys Lys Lys Lys Thr Asp Glu Ala Gln Pro Leu Pro
370 375 380

Gln Arg Gln Lys Lys Gln Pro Thr Val Thr Leu Leu Pro Ala Ala Asp
385 390 395 400

Met Asp Asp Phe Ser Arg Gln Leu Gln Asn Ser Met Ser Gly Ala Ser
405 410 415

Ala Asp Ser Thr Gln Ala
420

<210> 38
<211> 1377
<212> ADN
<213> CORONAVIRUS

<400> 38

atgaagggtca ccaaactgct gcatttagag acgtacttgt tgttttaa	aaacgaacaa	60
attaaaatgt ctgataatgg accccaatca aaccaacgta gtgcccccg	cattacattt	120
ggtggacca cagattcaac tgacaataac cagaatggag gacgcaatgg	ggcaaggcca	180
aaacagcgcc gacccaagg tttaccaat aatactgcgt cttgggtcac	agctctcact	240
cagcatggca aggaggaact tagattccct cgaggccagg gcgttccaat	caacaccaat	300
agtgggtccag atgaccaa	at tggctactac cgaagagcta cccgacgagt tcgtgggtggt	360
gacggcaaaa tgaaagagct cagccccaga tggtagttct attacctagg	aactggccca	420
gaagcttcac ttccctacgg cgctaacaaa gaaggcatcg tatgggttgc	aactgagggga	480
gccttgaata caccacaaaga ccacattggc acccgcaatc ctaataacaa	tgctgccacc	540
gtgctacaac ttcttcaagg aacaacattg ccaaaaggct tctacgcaga	gggaagcaga	600
ggcggcagtc aagcctcttc tcgctcctca tcacgtagtc gcggtaatc	aagaaattca	660
actcctggca gcagtagggg aaattctcct gctcgaatgg ctagcggagg	tggtgaaact	720
gccctcgcgc tattgctgct agacagattg aaccagcttg agagcaaagt	ttctggtaaa	780
ggccaacaac aacaaggcca aactgtcact aagaaatctg ctgctgaggc	atctaaaaag	840
cctcgccaaa aacgtactgc cacaaaacag tacaacgtca ctcaagcatt	tgaggagacgt	900

ggtccagaac aaaccaag	aaatttcggg gaccaagacc	taatcagaca aggaactgat	960
tacaaacatt ggccgcaaat	tgcacaattt gctccaagtg	cctctgcatt ctttggaatg	1020
tcacgcattg gcatggaagt	cacaccttcg ggaacatggc	tgacttatca tggagccatt	1080
aaattggatg acaaagatcc	acaattcaaa gacaacgtca	tactgctgaa caagcacatt	1140
gacgcataca aaacattccc	accaacagag cctaaaaagg	acaaaaagaa aaagactgat	1200
gaagctcagc ctttgccgca	gagacaaaag aagcagccca	ctgtgactct tcttcctgcg	1260
gctgacatgg atgattttctc	cagacaactt caaaattcca	tgagtggagc ttctgctgat	1320
tcaactcagg cataaacact	catgatgacc acacaaggca	gatgggctat gtaaacg	1377

<210> 39
 <211> 204
 <212> ADN
 <213> CORONAVIRUS

<400> 39			
atattaggtt tttacctacc	caggaaaagc caaccaacct	cgatctcttg tagatctgtt	60
ctctaaacga actttaaaat	ctgtgtagct gtcgctcggc	tgcatgccta gtgcacctac	120
gcagtataaa caataataaa	ttttactgtc gttgacaaga	aacgagtaac tcgtccctct	180
tctgcagact gcttacgggt	tcgt		204

<210> 40
 <211> 809
 <212> ADN
 <213> CORONAVIRUS

<400> 40			
actcaagcat ttgggagacg	tggtccagaa caaaccgaag	gaaatttcgg ggaccaagac	60
ctaatacagac aaggaactga	ttacaaacat tggccgcaaa	ttgcacaatt tgctccaagt	120
gcctctgcat tctttggaat	gtcacgcatt ggcatggaag	tcacaccttc gggaacatgg	180
ctgacttatc atggagccat	taaattggat gacaaagatc	cacaattcaa agacaacgtc	240
atactgctga acaagcacat	tgacgcatac aaaacattcc	caccaacaga gcctaaaaag	300
gacaaaaaga aaaagactga	tgaagctcag cttttgccgc	agagacaaaa gaagcagccc	360
actgtgactc ttcttcctgc	ggctgacatg gatgatttct	ccagacaact tcaaaattcc	420
atgagtggag cttctgctga	ttcaactcag gcataaacac	tcattgatgac cacacaaggc	480
agatgggcta tgtaaactgt	ttcgcaattc cgtttacgat	acatagtcta ctcttgtgca	540
gaatgaattc tcgtaactaa	acagcacaaag taggtttagt	taactttaat ctcacatagc	600
aatctttaat caatgtgtaa	cattagggag gacttgaaag	agccaccaca ttttcacgca	660
ggccacgcgg agtacgatcg	agggtagagt gaataatgct	aggagagct gcctatatgg	720
aagagcccta atgtgtaaaa	ttaatttttag tagtgctatc	cccatgtgat tttaatagct	780
tcttaggaga atgacaaaaa	aaaaaaaaa		809

<210> 41

<211> 448
<212> ADN
<213> CORONAVIRUS

<400> 41

aatgaacaca tagggctggt caagctgggg cagtacgcct tttccagct ctactagacc	60
acaagtgccca tttttgaggt gttcacgtgc ctccgatagg gcctcttcca cagagtcccc	120
gaagccacgc actagcacgt ctctaacctg aaggacaggc aaactgagtt ggacgtgtgt	180
tttctcggtg acaccaagaa caaggctctc catcttacct ttcggtcaca cccggacgaa	240
acctaggtat gctgatgatc gactgcaaca cggacgaaac cgtaagcagt ctgcagaaga	300
gggacgagtt actcgtttct tgtcaacgac agtaaaattt attattgttt atactgcgta	360
ggtgcactag gcatgcagcc gagcgacagc tacacagatt ttaaagttcg tttagagaac	420
agatctacaa gagatcgagg ttggttgg	448

<210> 42
<211> 2033
<212> ADN
<213> CORONAVIRUS

<400> 42

atacctaggt ttcgtccggg tgtgaccgaa aggtaagatg gagagccttg ttcttggtgt	60
caacgagaaa acacacgtcc aactcagttt gcctgtcctt caggttagag acgtgctagt	120
gcgtggcttc ggggactctg tggaagaggc cctatcggag gcacgtgaaC acctcaaaaa	180
tggcacttgt ggtctagtag agctggaaaa aggcgtactg cccagcttg aacagcccta	240
tgtgttcatt aaacgttctg atgccttaag caccaatcac ggccacaagg tcgttgagct	300
ggttgacagaa atggacggca ttcagtacgg tcgtagcggg ataacactgg gagtactcgt	360
gccacatgtg ggcgaaaccc caattgcata ccgcaatgtt cttcttcgta agaacggtaa	420
taagggagcc ggtggtcata gctatggcat cgatctaaag tcttatgact taggtgacga	480
gcttggcact gatcccattg aagattatga acaaaactgg aacactaagC atggcagtg	540
tgcactccgt gaactcactc gtgagctcaa tggaggtgca gtcactcgct atgtcgacaa	600

caatttctgt	ggcccagatg	ggtaccctct	tgattgcatc	aaagattttc	tcgcacgcgc	660
gggcaagtca	atgtgcactc	tttccgaaca	acttgattac	atcgagtcga	agagaggtgt	720
ctactgctgc	cgtgaccatg	agcatgaaat	tgcttggttc	actgagcgct	ctgataagag	780
ctacgagcac	cagacaccct	tcgaaattaa	gagtgccaaag	aaatttgaca	ctttcaaagg	840
ggaatgccca	aagtttgtgt	ttcctcttaa	ctcaaaagtc	aaagtcattc	aaccacgtgt	900
tgaaaagaaa	aagactgagg	gtttcatggg	gcgtatacgc	tctgtgtacc	ctgttgcac	960
tccacaggag	tgtaacaata	tgcaattgtc	taccttgatg	aaatgtaatc	attgcatga	1020
agtttcatgg	cagacgtgcg	actttctgaa	agccacttgt	gaacattgtg	gcactgaaaa	1080
tttagttatt	gaaggaccta	ctacatgtgg	gtacctacct	actaatgctg	tagtgaaaat	1140
gccatgtcct	gcctgtcaag	accagagat	tggacctgag	catagtgttg	cagattatca	1200
caaccactca	aacattgaaa	ctcgactccg	caaggagggt	aggactagat	gttttggagg	1260
ctgtgtgttt	gcctatgttg	gctgctataa	taagcgtgcc	tactgggttc	ctcgtgctag	1320
tgctgatatt	ggctcaggcc	atactggcat	tactggtgac	aatgtggaga	ccttgaatga	1380
ggatctcctt	gagatactga	gtcgtgaacg	tgtaaacatt	aacattgttg	gcgattttca	1440
tttgaatgaa	gaggttgcc	tcattttggc	atctttctct	gcttctacaa	gtgcctttat	1500
tgacactata	aagagtcttg	attacaagtc	tttcaaaacc	attgttgagt	cctgcggtaa	1560
ctataaagtt	accaagggaa	agcccgtaaa	agggtgcttg	aacattggac	aacagagatc	1620
agttttaaca	ccactgtgtg	gttttccctc	acaggctgct	ggtgttatca	gatcaatttt	1680
tgcgcgaca	cttgatgcag	caaaccactc	aattcctgat	ttgcaaagag	cagctgtcac	1740
catacttgat	ggtattttctg	aacagtcatt	acgtcttgtc	gacgccatgg	tttatacttc	1800
agacctgctc	accaacagtg	tcattattat	ggcatatgta	actggtgggtc	ttgtacaaca	1860
gactttctcag	tggttgtcta	atcttttggg	cactactgtt	gaaaaactca	ggcctatctt	1920
tgaatggatt	gaggcgaaac	ttagtgcagg	agttgaattt	ctcaaggatg	cttgggagat	1980
tctcaaattt	ctcattacag	gtgtttttga	catcgtcaag	ggtcaaatac	agg	2033

<210> 43
 <211> 2018
 <212> ADN
 <213> CORONAVIRUS

<400> 43

ggattgaggc	gaaacttagt	gcaggagttg	aattttctcaa	ggatgcttgg	gagattctca	60
aattttctcat	tacaggtgtt	tttgacatcg	tcaaggggtca	aatacagggtt	gcttcagata	120
acatcaagga	ttgtgtaaaa	tgcttcattg	atgttggttaa	caaggcactc	gaaatgtgca	180
ttgatcaagt	cactatcgct	ggcgcaaagt	tgcatcact	caacttaggt	gaagtcttca	240

tcgctcaaag caagggactt taccgtcagt gtatacgtgg caaggagcag ctgcaactac	300
tcatgcctct taaggcacca aaagaagtaa cctttcttga aggtgattca catgacacag	360
tacttacctc tgaggaggtt gttctcaaga acggtgaact cgaagcactc gagacgcccg	420
ttgatagctt cacaaatgga gctatcgttg gcacaccagt ctgtgtaaata ggcctcatgc	480
tcttagagat taaggacaaa gaacaatact gcgcattgtc tcctggttta ctggctacaa	540
acaatgtctt tcgcttaaaa ggggggtgcac caattaaagg tgtaaccttt ggagaagata	600
ctgtttggga agttcaaggt tacaagaatg tgagaatcac atttgagctt gatgaacgtg	660
ttgacaaagt gcttaatgaa aagtgtcttg tctacactgt tgaatccggt accgaagtta	720
ctgagtttgc atgtgttgta gcagaggctg ttgtgaagac ttacaacca gtttctgata	780
tccttaccaa catgggtatt gatcttgatg agtggagtgt agctacattc tacttatttg	840
atgatgctgg tgaagaaaac ttttcatcac gtatgtattg ttctttttac cctccagatg	900
aggaagaaga ggacgatgca gagtgtgagg aagaagaaat tgatgaaacc tgtgaacatg	960
agtacggtac agaggatgat tatcaaggct tccctctgga atttggtgcc tcagctgaaa	1020
cagttcgagt tgaggaagaa gaagaggaag actggctgga tgatactact gagcaatcag	1080
agattgagcc agaaccagaa cctacacctg aagaaccagt taatcagttt actggttatt	1140
taaaacttac tgacaatgtt gccattaat gtgttgacat cgtaaggag gcaaaaagtg	1200
ctaactctat ggtgattgta aatgctgcta acatacacct gaaacatggt ggtggtgtag	1260
caggtgcact caacaaggca accaatggtg ccatgcaaaa ggagagtgat gattacatta	1320
agctaaatgg cctctttaca gtaggagggg cttgtttgct ttctggacat aatcttgcta	1380
agaagtgtct gcatgttggt ggacctaac taaatgcagg tgaggacatc cagcttctta	1440
aggcagcata tgaataattc aattcacagg acatcttact tgcaccattg ttgtcagcag	1500
gcatatttgg tgctaaacca cttcagtcct tacaagtgtg cgtgcagacg gttcgtacac	1560
aggtttatat tgcagtcaat gacaaagctc tttatgagca ggttgatcat gattatcttg	1620
ataacctgaa gcctagagtg gaagcaccta aacaagagga gccaccaaac acagaagatt	1680
ccaaaactga ggagaaatct gtcgtacaga agcctgtcga tgtgaagcca aaaattaagg	1740
cctgcattga tgaggttacc acaacactgg aagaaactaa gtttcttacc aataagttac	1800
tcttgtttgc tgatatcaat ggtaagcttt accatgattc tcagaacatg cttagaggtg	1860
aagatatgtc tttccttgag aaggatgcac cttacatggt aggtgatgtt atcactagt	1920
gtgatatac ttgtgttgta ataccctcca aaaaggctgg tggcactact gagatgctct	1980
caagagcttt gaagaaagtg ccagttgatg agtatata	2018

<210> 44
 <211> 1442
 <212> ADN
 <213> CORONAVIRUS

 <400> 44

ttgatgaggt taccacaaca ctggaagaaa ctaagtttct taccaataag ttactcttgt	60
ttgctgatat caatggtaag ctttaccatg attctcagaa catgcttaga ggtgaagata	120
tgtcttttct tgagaaggat gcaccttaca tggtaggtga tgttatcact agtggtgata	180
tcacttgtgt tgtaataccc tccaaaaagg ctggtggcac tactgagatg ctctcaagag	240
ctttgaagaa agtgccagtt gatgagtata taaccacgta ccctggacaa ggatgtgctg	300
gttatacact tgaggaagct aagactgctc ttaagaaatg caaatctgca ttttatgtac	360
taccttcaga agcacctaata gctaaggaag agattctagg aactgtatcc tggaatttga	420
gagaaatgct tgtcatgct gaagagacaa gaaaattaat gcctatatgc atggatgtta	480
gagccataat ggcaaccatc caacgtaagt ataaaggaat taaaattcaa gagggcatcg	540
ttgactatgg tgtccgattc ttcttttata ctagtaaaga gcctgtagct tctattatta	600
cgaagctgaa ctctctaaat gagccgcttg tcacaatgcc aattgggttat gtgacacatg	660
gttttaatat tgaagaggct gcgcgctgta tgcgttctct taaagctcct gccgtagtgt	720
cagtatcatc accagatgct gttactacat ataattgata cctcacttcg tcatcaaaga	780
catctgagga gcactttgta gaaacagttt ctttggctgg ctcttacaga gattggctct	840
attcaggaca gcgtacagag ttaggtgttg aatttcttaa gcgtggtgac aaaattgtgt	900
accacactct ggagagcccc gtcgagtttc atcttgacgg tgaggttctt tcacttgaca	960
aactaaagag tctcttatcc ctgcgggagg ttaagactat aaaagtgttc acaactgtgg	1020
acaacactaa tctccacaca cagcttgtgg atatgtctat gacatatgga cagcagtttg	1080
gtccaacata cttggatggg gctgatgtta caaaaattaa acctcatgta aatcatgagg	1140
gtaagacttt ctttgtacta cctagtgatg acacactacg tagtgaagct ttcgagtact	1200
accatactct tgatgagagt ttcttggta ggtacatgtc tgctttaaac cacacaaaga	1260
aatggaaatt tcctcaagtt ggtggtttaa cttcaattaa atgggctgat aacaattggt	1320
atttgtctag tgttttatta gcacttcaac agcttgaagt caaattcaat gcaccagcac	1380
ttcaagaggc ttattataga gcccgtgctg gtgatgctgc taacttttgt gcactcatac	1440
tc	1442

<210> 45
 <211> 1050
 <212> ADN
 <213> CORONAVIRUS

<400> 45	
atatgtctat gacatatgga cagcagtttg gtccaacata cttggatggg gctgatgtta	60

caaaaattaa acctcatgta aatcatgagg gtaagacttt ctttgtacta cctagtgatg	120
acacactacg tagtgaagct ttcgagtact accatactct tgatgagagt tttcttggtg	180
ggtacatgtc tgctttaaac cacacaaaga aatggaaatt tcctcaagtt ggtggtttaa	240
cttcaattaa atgggctgat aacaattggt atttgtctag tgttttatta gcacttcaac	300
agcttgaagt caaattcaat gcaccagcac ttcaagaggc ttattataga gcccgtgctg	360
gtgatgctgc taacttttgt gcactcatac tcgcttacag taataaaact gttggcgagc	420
ttggtgatgt cagagaaact atgacccatc ttctacagca tgctaatttg gaatctgcaa	480
agcgagttct taatgtggtg tgtaaacatt gtggtcagaa aactactacc ttaacgggtg	540
tagaagctgt gatgtatatg ggtactctat cttatgataa tcttaagaca ggtgtttcca	600
ttccatgtgt gtgtggctgt gatgctacac aatatctagt acaacaagag tcttcttttg	660
ttatgatgtc tgcaccacct gctgagtata aattacagca aggtacattc ttatgtgcga	720
atgagtacac tggtaactat cagtgtggtc attacactca tataactgct aaggagaccc	780
tctatcgtat tgacggagct caccttaca agatgtcaga gtacaaagga ccagtgactg	840
atgttttcta caaggaaaca tcttacctca caaccatcaa gcctgtgtcg tataaactcg	900
atggagttac ttacacagag attgaaccaa aattggatgg gtattataaa aaggataatg	960
cttactatac agagcagcct atagaccttg taccaactca accattacca aatgcgagtt	1020
ttgataattt caaactcaca tgttctaaca	1050

<210> 46
 <211> 1995
 <212> ADN
 <213> CORONAVIRUS

<400> 46

tttgtgcact catactcgct tacagtaata aaactgttgg cgagcttggt gatgtcagag	60
aaactatgac ccatcttcta cagcatgcta atttggaatc tgcaaagcga gttcttaatg	120
tggtgtgtaa acattgtggt cagaaaacta ctaccttaac ggggtgtagaa gctgtgatgt	180
atatgggtac tctatcttat gataatctta agacaggtgt ttccattcca tgtgtgtgtg	240
gtcgtgatgc tacacaatat ctagtacaac aagagtcctc ttttgttatg atgtctgcac	300
cacctgctga gtataaatta cagcaaggta cattcttatg tgcgaatgag tacactggta	360
actatcagtg tggtcattac actcatataa ctgctaagga gacctctat cgtattgacg	420
gagctcacct tacaagatg tcagagtaca aaggaccagt gactgatgtt ttctacaagg	480
aaacatctta cactacaacc atcaagcctg tgtcgtataa actcgatgga gttacttaca	540
cagagattga accaaaattg gatgggtatt ataaaaagga taatgcttac tatacagagc	600
agcctataga cttgtacca actcaaccat taccaaatgc gagttttgat aatttcaaac	660

tcacatgttc	taacacaaaa	tttgctgatg	atttaaataca	aatgacaggc	ttcacaaagc	720
cagcttcacg	agagctatct	gtcacattct	tcccagactt	gaatggcgat	gtagtggcta	780
ttgactatag	acactattca	gcgagtttca	agaaagggtgc	taaattactg	cataagccaa	840
ttgtttggca	cattaaccag	gctacaacca	agacaacgtt	caaaccacaa	acttggtggt	900
tacgttgtct	ttggagtaca	aagccagtag	atacttcaaa	ttcatttgaa	gttctggcag	960
tagaagacac	acaaggaatg	gacaatcttg	cttgtgaaag	tcaacaaccc	acctctgaag	1020
aagtagtgga	aaatcctacc	atacagaagg	aagtcataga	gtgtgacgtg	aaaactaccg	1080
aagttgtagg	caatgtcata	cttaaaccat	cagatgaagg	tgttaaagta	acacaagagt	1140
taggtcatga	ggatcttatg	gctgcttatg	tggaaaacac	aagcattacc	attaagaaac	1200
ctaagtagct	ttcactagcc	ttaggtttaa	aaacaattgc	cactcatggt	attgctgcaa	1260
ttaatagtgt	tccttggagt	aaaatttttg	cttatgtcaa	accattctta	ggacaagcag	1320
caattacaac	atcaaattgc	gctaagagat	tagcacaacg	tgtgtttaac	aattatatgc	1380
cttatgtgtt	tacattattg	ttccaattgt	gtacttttac	taaaagtacc	aattctagaa	1440
ttagagcttc	actacctaca	actattgcta	aaaatagtgt	taagagtgtt	gctaaattat	1500
gtttggatgc	cggcattaat	tatgtgaagt	cacccaaatt	ttctaaattg	ttcacaatcg	1560
ctatgtggct	attgtttgta	agtatttgct	taggttctct	aatctgtgta	actgctgctt	1620
ttggtgtact	cttatctaata	tttggtgctc	cttcttattg	taatggcggt	agagaattgt	1680
atcttaattc	gtctaacgtt	actactatgg	atttctgtga	aggttctttt	ccttgcagca	1740
tttgtttaag	tggattagac	tcccttgatt	cttatccagc	tcttgaaacc	attcaggtga	1800
cgatttcac	gtacaagcta	gacttgacaa	tttttaggtct	ggccgctgag	tgggttttgg	1860
catatatgtt	gttcacaaaa	ttcttttatt	tattaggtct	ttcagctata	atgcaggtgt	1920
tctttggcta	ttttgctagt	catttcatca	gcaattcttg	gctcatgtgg	tttatcatta	1980
gtattgtaca	aatgg					1995

<210> 47
 <211> 1884
 <212> ADN
 <213> CORONAVIRUS

aattcttggc	tcatgtgggt	tatcattagt	attgtacaaa	tggcaccggt	ttctgcaatg	60
gttaggatgt	acatcttctt	tgcttctttc	tactacatat	ggaagagcta	tggtcatatc	120
atggatgggt	gcacctcttc	gacttgcatg	atgtgctata	agcgcaatcg	tgccacacgc	180
gttgagtgtg	caactattgt	taatggcatg	aagagatctt	tctatgtcta	tgcaaatgga	240
ggccgtggct	tctgcaagac	tcacaattgg	aattgtctca	attgtgacac	attttgcact	300

ggtagtacat tcattagtgat tgaagttgct cgtgatttgt cactccagtt taaaagacca	360
atcaacccta ctgaccagtc atcgtatatatt gttgatagtg ttgctgtgaa aaatggcgcg	420
cttcacctct acttttgacaa ggctgggtcaa aagacctatg agagacatcc gctctcccat	480
tttgtcaatt tagacaattt gagagctaac aacactaaag gttcactgcc tattaatgtc	540
atagtttttg atggcaagtc caaatgcgac gagtctgctt ctaagtctgc ttctgtgtac	600
tacagtcagc tgatgtgcca acctattctg ttgcttgacc aagctcttgt atcagacggt	660
ggagatagta ctgaagtttc cgtaagatg tttgatgctt atgtcgacac cttttcagca	720
acttttagtg ttcttatgga aaaacttaag gcacttggtg ctacagctca cagcgagtta	780
gcaaaggggtg tagcttttaga tgggtgtcctt tctacattcg tgtcagctgc ccgacaagggt	840
gttggttgata ccgatgttga cacaaggat gttattgaat gtctcaaact ttcacatcac	900
tctgacttag aagtgcagag tgacagttgt aacaatttca tgctcaccta taataagggt	960
gaaaacatga cgcccagaga tcttgggcgca tgtattgact gtaatgcaag gcatatcaat	1020
gcccagtag caaaaagtca caatgtttca ctcatctgga atgtaaaaga ctacatgtct	1080
ttatctgaac agctgcgtaa acaaattcgt agtgctgcca agaagaacaa catacctttt	1140
agactaactt gtgctacaac tagacaggtt gtcaatgtca taactactaa aatctcactc	1200
aagggtggtg agattgtag tacttgtttt aaacttatgc ttaaggccac attattgtgc	1260
gttcttgctg cattgggttg ttatatcgtt atgccagtag atacattgtc aatccatgat	1320
ggttacacaa atgaaatcat tggttacaaa gccattcagg atgggtgtcac tcgtgacatc	1380
atttctactg atgattgttt tgcaaataaa catgctgggt ttgacgcatg gtttagccag	1440
cgtgggtggtt catacaaaaa tgacaaaagc tgccctgtag tagctgctat cattacaaga	1500
gagattggtt tcatagtgcc tggcttaccg ggtactgtgc tgagagcaat caatggtgac	1560
ttcttgcat tttctacctg tgtttttagt gctgttgcca acatttgcta cacaccttcc	1620
aaactcattg agtatagtga ttttgctacc tctgcttgcg ttcttgctgc tgagtgtaca	1680
atttttaagg atgctatggg caaacctgtg ccatattgtt atgacactaa tttgctagag	1740
ggttctatct cttatagtga gcttcgtcca gacactcggt atgtgcttat ggatgggtcc	1800
atcatacagt ttcttaacac ttacctggag ggttctgtta gagtagtaac aacttttgat	1860
gctgagtact gtagacatgg taca	1884

<210> 48
 <211> 2020
 <212> ADN
 <213> CORONAVIRUS

<400> 48

cactcgttat gtgcttatgg atggttccat catacagttt cctaacactt acctggaggg	60
---	----

ttctgttaga gtagtaacaa cttttgatgc tgagtactgt agacatggta catgcgaaag	120
gtcagaagta ggtatTTGCC tatctaccag tggtagatgg gttcttTaata atgagcatta	180
cagagctcta tcaggagttt tctgtggtgt tgatgcgatg aatctcatag ctaacatctt	240
tactcctctt gtgcaacctg tgggtgcttt agatgtgtct gcttcagtag tggctgggtg	300
tattattgcc atattggtga cttgtgctgc ctactacttt atgaaattca gacgtgtttt	360
tgggtgagtac aaccatgttg ttgctgctaa tgcacttttg tttttgatgt ctttactat	420
actctgtctg gtaccagctt acagctttct gccgggagtc tactcagtct tttacttgta	480
cttgacattc tatttcacca atgatgtttc attcttggct caccttcaat ggtttgccat	540
gttttctcct attgtgcctt tttggataac agcaatctat gtattctgta tttctctgaa	600
gcactgccat tggttcttta acaactatct taggaaaaga gtcattgtta atggagttac	660
atttagtacc ttcgaggagg ctgctttgtg tacctttttg ctcaacaagg aaatgtacct	720
aaaattgcgt agcgagacac tgttgccact tacacagtat aacaggatatc ttgctctata	780
taacaagtac aagtatttca gtggagcctt agatactacc agctatcgtg aagcagcttg	840
ctgccactta gcaaaggctc taaatgactt tagcaactca ggtgctgatg ttctctacca	900
accaccacag acatcaatca cttctgctgt tctgcagagt ggttttagga aaatggcatt	960
cccgtcaggc aaagttgaag ggtgcatggt acaagtaacc tgtggaacta caactcttaa	1020
tggattgtgg ttggatgaca cagtatactg tccaagacat gtcatttgca cagcagaaga	1080
catgcttaat cctaactatg aagatctgct cattcgcaaa tccaaccata gctttcttgt	1140
tcaggctggc aatgttcaac ttcgtgttat tggccattct atgcaaaatt gtctgcttag	1200
gcttaaagtt gatacttcta accctaagac acccaagtat aaatttgtcc gtatccaacc	1260
tgggtcaaaca ttttcagttc tagcatgcta caatggttca ccatctgggtg tttatcagtg	1320
tgccatgaga cctaatacata ccattaaagg ttctttcctt aatggatcat gtggtagtgt	1380
tggttttaac attgattatg attgctgtgc tttctgctat atgcatcata tggagcttcc	1440
aacaggagta cacgctggta ctgacttaga aggtaaattc tatgggtccat ttgttgacag	1500
acaaactgca caggctgcag gtacagacac aaccataaca ttaaatgttt tggcatggct	1560
gtatgctgct gttatcaatg gtgatagggtg gtttcttaat agattcacca ctactttgaa	1620
tgactttaac cttgtggcaa tgaagtacaa ctatgaacct ttgacacaag atcatgttga	1680
catattggga cctctttctg ctcaaacagg aattgccgtc ttagatatgt gtgctgcttt	1740
gaaagagctg ctgcagaatg gtatgaatgg tcgtactatc cttggtagca ctattttaga	1800
agatgagttt acaccatttg atgttgttag acaatgctct ggtgttacct tccaaggtaa	1860
gttcaagaaa attgttaagg gcactcatca ttggatgctt ttaactttct tgacatcact	1920
attgattctt gttcaaagta cacagtgggc actgtttttc tttgtttacg agaatgcttt	1980
cttgccattt actcttggtg ttatggcaat tgctgcatgt	2020

<211> 2040
<212> ADN
<213> CORONAVIRUS

<400> 49

agcatttcca	gcctgaagac	gtactgtagc	agctaaactg	cccagcacca	tacctctatt	60
taggttggtt	aagcctttga	tgaagtacaa	gtatttcact	ttaggccctt	ttggtgtgtc	120
tgtaacaaac	ctacaagggtg	gttccaggtc	tgtgtaaatt	gtacctgtac	catcactctt	180
agggaaatcta	gcccatttga	gatcttggtg	gtctgatagt	aatgccagca	caaacctacc	240
tcccttcgaa	ttgttatagt	aggcaagtgc	attgtcatca	gtacaagctg	tttgtgtggt	300
accagccgca	caggacatct	gtcgtagtgc	tactggactc	agttcattat	tctgtagttt	360
aacagctgag	ttggctctta	gagctgtaac	aataagaggc	caagccaaat	ttggtgaatt	420
gtccatgtta	atttactaa	gttgaacaat	cttgcctatcc	gcatcaacaa	cttgctggat	480
ttcccagagt	gcagatgcat	atgtaaaggt	gttaccatca	caagtgttct	tgtagggtacc	540
ataatcaggg	acaacaacca	tgagtttggc	tgctgtagtc	aatggatatga	tgttgagtgg	600
aacacaacca	tcacgcgcat	tgttgataat	gttgtaagt	gcatcattat	caagcttcct	660
aagcatagtg	aagagcattg	tttgcatagc	actagttact	tttgccctct	tgtcctcaga	720
tcttgccctgt	ttgtacattt	gggtcatagc	ctgatctgcc	atcttttcca	acttgcggtg	780
catggcagca	tcacggtcaa	actcagattt	agccacattc	aaagatttct	ttaacttttt	840
gagaacgact	tcagaatcac	cattagctac	agcctgctca	taggcctcct	gggcagtggc	900
ataagcggca	tatgatggta	aagaactaaa	ttctgaagca	atagcctgaa	gagtagcacg	960
gttatcgagc	atttcctcgc	acaacctatt	aatgtctaca	gcaccctgca	tggatagcaa	1020
aacagacaaa	agagaaacca	tcttctcgaa	agcttcagtt	gtgtcttttg	caagaagaat	1080
atcattgttg	agttgtacac	attgtgcca	caatttagaa	gatgactcta	ctctaagttg	1140
ttgaagaacc	gagagcagta	ccacagatgt	gcactttacg	tcagacattt	tagactgtac	1200
agtagcaacc	ttgatacatg	gtttacctcc	aatacccaac	aacttaatgt	taagcttgaa	1260
agcatcaata	ctactcttag	gaggcaaaag	cccctgggag	ttcatatacc	taaattcttg	1320
tgtagagacc	aagtagtcat	aaacaccaag	agtaagcctg	aagtaacggg	tgagtaaaca	1380
gaaaaggcca	aagtagcagc	agcaacaata	gcctaagaaa	caataaacia	gcatgataca	1440
ctgtaagggtg	ttgccagtaa	taaataacia	tggttaatac	tcaacacaca	caaacactat	1500
agctctagct	aaaaacatga	tagtcgtaac	gacaccagaa	tagttagagg	ttacagaaat	1560
aactaaggcc	cacatggaaa	tagcttgatc	taaagcatta	ccatagtaga	ctttgtaaac	1620
aagtgtaatg	acattcatca	gtgtccaaac	acgtctagca	gcatcatcat	aaacagtgcg	1680
agctgtcatg	agaataagca	aaactaaagc	tgaagcatac	ataacaCaat	ccttaagcct	1740

ataaccagac	aagctagtgt	cagccaattc	aagccatgtc	atgatacgca	tcacccagct	1800
agcaggcatg	tagaccatat	taaagtaagc	aactgttgca	agagaaggta	acagaaacaa	1860
gcacaagaat	gcgtgcttat	gcttaacaag	cagcatagca	catgcagcaa	ttgccataat	1920
accaagagta	aatggcaaga	aagcattctc	gtaaacaaag	aaaaacagtg	accactgtgt	1980
actttgaaca	agaatcaata	gtgatgtcaa	gaaagttaaa	agcatccaat	gatgagtgc	2040

<210> 50

<211> 2012

<212> ADN

<213> CORONAVIRUS

<400> 50

cttgtagggt	tgttacagac	acaccaaaaag	ggcctaaagt	gaaataacttg	tacttcatca	60
aaggcttaaa	caacctaaat	agaggatatg	tgctgggcag	tttagctgct	acagtacgtc	120
ttcaggctgg	aaatgctaca	gaagtacctg	ccaattcaac	tgtgctttcc	ttctgtgctt	180
ttgcagtaga	ccctgctaaa	gcatataagg	attacctagc	aagtggagga	caaccaatca	240
ccaactgtgt	gaagatgttg	tgtacacaca	ctggtagagg	acaggcaatt	actgtaacac	300
cagaagctaa	catggacca	gagtcctttg	gtgggtgctt	atgttgtctg	tattgtagat	360
gccacattga	ccatccaaat	cctaaaggat	tctgtgactt	gaaaggtaag	tacgtccaaa	420
tacctaccac	ttgtgcta	gaccagtg	gttttacact	tagaaacaca	gtctgtaccg	480
tctgcggaat	gtggaaagg	tatggctgta	gttgtagacca	actccgcgaa	cccttgatgc	540
agtctgcgga	tgcatcaacg	tttttaa	ggtttgcggt	gtaagtgcag	cccgtcttac	600
accgtgcggc	acaggcacta	gtactgatgt	cgtctacagg	gcttttgata	tttacaacga	660
aaaagttgct	ggttttgcaa	agttcctaaa	aactaattgc	tgctgcttcc	aggagaagga	720
tgaggaaggc	aattttattag	actcttactt	tgtagttaag	aggcatacta	tgtctaacta	780
ccaacatgaa	gagactat	ataacttgg	taaagattgt	ccagcggttg	ctgtccatga	840
ctttttcaag	tttagagtag	atggtgacat	ggtaccacat	atatcacgtc	agcgtcta	900
taaatacaca	atggctgatt	tagtctatgc	tctacgtcat	tttgatgagg	gtaattgtga	960
tacattaaaa	gaaataactg	tcacatacaa	ttgctgtgat	gatgattatt	tcaataagaa	1020
ggattggtat	gacttcgtag	agaatcctga	catcttacgc	gtatatgcta	acttaggtga	1080
gcgtgtacgc	caatcattat	taaagactgt	acaattctgc	gatgctatgc	gtgatgcagg	1140
cattgtaggc	gtactgacat	tagataatca	ggatctta	gggaactgg	acgatttcgg	1200
tgatttcgta	caagtagcac	caggctgcgg	agttcctatt	gtggattcat	attactcatt	1260
gctgatgcc	atcctcactt	tgactagggc	attggctgct	gagtcacata	tggatgctga	1320
tctcgcaaaa	ccacttatta	agtgggattt	gctgaaatat	gattttacgg	aagagagact	1380

```

ttgtctcttc gaccgttatt ttaaataattg ggaccagaca taccatccca attgtattaa 1440
ctgtttggat gataggtgta tccttcattg tgcaaaacttt aatgtgttat tttctactgt 1500
gtttccacct acaagttttg gaccactagt aagaaaaata tttgtagatg gtgttccttt 1560
tgttgtttca actggatacc attttcgtga gttaggagtc gtacataatc aggatgtaaa 1620
cttacatagc tcgcgtctca gtttcaagga acttttagtg tatgctgctg atccagctat 1680
gcatgcagct tctggcaatt tattgctaga taaacgcact acatgctttt cagtagctgc 1740
actaacaac aatgttgctt ttcaaactgt caaacccggt aattttaata aagactttta 1800
tgactttgct gtgtctaaag gtttctttaa ggaaggaagt tctgttgaac taaaacactt 1860
cttctttgct caggatggca acgctgctat cagtgattat gactattatc gttataatct 1920
gccaacaatg tgtgatatca gacaactcct attcgtagtt gaagttgttg ataaataactt 1980
tgattgttac gatggtggct gtattaatgc ca 2012

```

<210> 51
 <211> 1877
 <212> ADN
 <213> CORONAVIRUS

<400> 51

```

gtacttcgcg tacagtggca ataccatatg acagcttaaa tgtttcctca gtggctttga 60
gcgtttctgc tgcgaaaagc ttgagtctct cagtacaagt gttggcaagt atgtaatcgc 120
cagcattagt ccaatcacat gttgctatcg cattgaagtc agtgacattg tcaactgccta 180
cacatgtgtt tttgtataaa ccaaaaacct gaccattagc acataatgga aaactaatgg 240
gaggcttatg tgacttgcaa taatagctca tacctcctag atacagttgt gtcacatcag 300
tgacatcaca acctggggca ttgcaaacat agggattaac agacaacact aatttgtgtg 360
atgttgaaat gacatggtca tagcagcact tgcaacatag gaatggtctc ctaatacagg 420
caccgcaacg aagtgaagtc tgtgaattgc acaatacaca agcacctaca gcctgcaaga 480
ctgtatgtgg tgtgtacata gcctcataaa actcaggttc ccagtaccgt gaggtgttat 540
cattagttag cattacggaa tacatgtcca acatgtggcc agtaagctca tcatgtaact 600
ttctaattga ttgtaaatac aagtgaagaa catcagcata ctcttgatta ggatgttttg 660
taagtgggta agcatcaata gccagtgaca cgaacctttc aatcataagt gtaccatctg 720
ttttgacaat atcatcgaca aaacagcctg cgcctaatat tcttgatgga tctgggtaag 780
gcaggtacac gtaatcatct ccttgtttaa ctagcattgt atgctgtgag caaaattcgt 840
gaggtccttt agtaagggtca gtctcagtcc aacattttgc ctcagacatg aacacattat 900
tttgataata aagaactgcc ttaaagttct taatgctagc tactaaacct tgagccgcat 960
agttactgtt atagcacaca acggcatcat cagaaagaat catcatggag aaatgtttac 1020

```

gcaggtaagc gtaaaactca tccacgaatt catgatcaac atccctatTT ctatagagac	1080
actcatagag cctgtgttgt agattgcgga catacttgtc agctatctta ttaccatcag	1140
ttgaaagaag tgcatttaca ttggctgtaa cagcttgaca aatgttaaag acactattag	1200
cataagcagt tgtagcatca ccggatgatg ttccacctgg tttaacatat agtgagccgc	1260
cacacatgac catctcactt aatacttgcg cacactcggt agctaacctg tagaaacggt	1320
gtgataagtt acagcaagtg ttatgtttgc gagcaagaac aagagaggcc attatcctaa	1380
gcatgttagg catggctctg tcacattttg gataatccca acccataagg tgtggagttt	1440
ctacatcact gtaaacagtt tttaacatat tatgccagcc accgtaaaac ttgcttggtc	1500
caattaccac agtagctcct ctagtggcgg ctattgactt caataatttc tgatgaaact	1560
gtctatttgt catagtacta cagatagaga caccagctac ggtgcgagct ctattcttg	1620
cactaatggc atacttaaga ttcatTTgag ttatagtagg gatgacatta cgcttagtat	1680
acgcgaaaag tgcattctga tcctcataac tcattgagtc ataataaagt ctagccttac	1740
cccatttatt aaatgggaaa ccagctgatt tatccagatt gttaacgatt acttggttg	1800
cattaataca gccaccatcg taacaatcaa agtatTTatc aacaacttca actacgaata	1860
ggagtTgtct gatatca	1877

<210> 52
 <211> 2051
 <212> ADN
 <213> CORONAVIRUS

<400> 52

tcaggTccaa tcttgacaaa gtacttcatt gatgtaagct caaagccatg cgcccaaagg	60
acgaacacga ctctgtctga caatcTTtc agtgatcac tgagcatttg tactatctta	120
atacgcacta cattccaggg caagcTTta tacatgagtg gtataagatg tttaaactgg	180
tcacctggtg gaggtTTtg attaaactctg gtgaattctg tgTTattttc agtgtcaaca	240
taaccagtcg gtacagctac taagttaaca cctgtagaaa atcctagctg gagaggtagg	300
ttagtaccga cagcatctct agttgcatga cagccctcta catcaaagcc aatccacgca	360
cgaacgtgac gaatagcttc ttcgcgggtg ataaacatat tagggtaacc attgacttg	420
taattcattt tgaaacccat catagagatg agtctacggg aggtcatgtc ctttggtatg	480
cctggtatgt caacacataa tccttcagtc ttgaacttta tatcaacgct gaggtgtgta	540
ggtgcctgtg taggatgaag accagtaatg atcttactac agtccttaaa aagtccagtt	600
acattttctg cttgtaattg agccacattg cgacgtggta tttctagact tgtaaattgc	660
agtttgtcat aaagatctct atcagacatt atgcacaaaa tgccaatttt tgcccttg	720
atagccacat tgaagcggtt gacattacaa gagtgtgctg tttcagtagt ttgtgtgaat	780

atgacatagt catattcaga accctgtgat gaatcaacag tctgcgtagg caatcctaag	840
atTTTTgaag ctacagcggt ctgtgaatta taagggtgaga taaaaacagc ttttctccaa	900
gcaggattgc gtgtaagaaa ttctcttaca acgcctatTT gaggtctggt gattgcagat	960
gaaacatcat gtgtaataac acctttgtag aacatTTTga agcattgagc tgacttatcc	1020
ttgtgtgctt ttagcttatt gtcataaact aaagcactca cagtgtcaac aatttcagca	1080
ggacaacggc gacaagttcc aaggaacatg tctggacctt ttgttttcat aagtctgcac	1140
actgaattaa aatattctgg ttctagtgtg cttttagtca gcaatgtgcg gggggctggt	1200
aattgagcag gatcgccaat atagacgtag tgttttgac gaagtctagc attgacaaca	1260
ctcaagtcatt aattagtagc catagagatt tcatcaaaga ctacaatgtc agcagttggt	1320
tctggcaatg catttacagt gcagaaaaca tactgttcta gtgttgaatt cactttgaat	1380
ttatcaaaac actctacgcg cgcacgcgca ggtatgattc tactacattt atctatgggc	1440
aaatattTTta atgccttttT acatagggca tcaacagctg catgagagca tgccgtatac	1500
actatgcgag cagatgggta atagagagca agtccgatgg caaaatgact cttaccagta	1560
ccagggtggc cttggagtggt agagtacttt tgcatgccga ctttttgata atttgcaaca	1620
ttgctagaaa actcatctga gatgttgagt gttgggtaca agccagtaat tctcacatag	1680
tgctcttggt gcactagagt aggtgcacta agtggcatta cagtgtgaga tgtcaacaca	1740
aagtaatcac caacattcaa cttgtatgtc gtagtacctc tgtacacaac agcatcacca	1800
tagtcacctt tttcaaagggt gtactctcca atctgtactt tactattttt agttacacgg	1860
taaccagtaa agacatagtt tctgttcaat ggtgggtctag gttttccaac ctcccatgaa	1920
agatgcaatt ctctgtcaga gagtacttcg cgtacagtggt caataccata tgacagctta	1980
aatgtttcct cagtggcttt gagcgtttct gctgcgaaaa gcttgagtct ctcagtacaa	2040
gtgttgga g	2051

<210> 53

<211> 2075

<212> ADN

<213> CORONAVIRUS

<400> 53

tgctttagt tttgggtaga aggtttcaac atgtccatcc ttacaccaa gcatgaatga	60
aatttcagca tagtcaattg taaccttgac cacttttgaa atcactgaca aatcttgga	120
ctttattatc tcgacaaagt catcaagtaa aagatcaatc acagaacaca cacattttga	180
tgaacctgtt tgcgcatctg ttatgaagta atttttcact gtgctgtcca tagggataaa	240
atcctctaatt ttaagtgggt aatcttgga gcgcttggt aagcctatca ttaaataag	300
accgccaagt tgtccatgac tgaaatctcc ataaacgatg tggtcgaagg catagccctc	360

gagcttatat	cgctgtatga	attcatccat	agcgagctcg	agaaagtcag	tttccatttg	420
tgatctgggc	ttaaaatcct	ctaagtctct	gctctgagta	aagtaggttt	caggcaactg	480
ttgaataatg	ccgtctactt	tcttaaagta	gttaaactgt	gtttttactg	attctccaat	540
taatgtgact	ccattgacgc	tagcttgtgc	tgggtccctt	gaagggtgta	gacctttgac	600
tgaaccttct	gttattaaaa	caccattacg	ggcgtttcta	aaaagggtcta	cctgtccttc	660
cactctacca	tcaaacaaga	cagtaagtga	agaacaagca	ctctcagtag	gtttcttggc	720
aatgtcagtc	attgtgcaga	cacctattgt	agatacatgt	gctggggctt	ctcttttgta	780
gtcccagatt	acagtattag	cagcgatata	aacacccaaa	ttattgagta	tcttaatctc	840
tggcactggg	ttaatgttac	gcttagccca	aagctcaaat	gcaacattaa	caggaagtgt	900
tgtcttattt	tcaaagatct	ccacatcaat	accatctacc	tttgtgtaaa	cagcattatt	960
aatgatggaa	acaggtgctt	cgccggcggt	tccatcaaag	tgtcctttat	taacaacatt	1020
ataagccaca	ttttctaaac	tctgtaacct	ggtaaatgta	ttccacaggt	tataagtatc	1080
aaattgtttg	taaatccata	ggctaaatcc	agcagaaatc	atcatattat	atgcatccaa	1140
gtactgtcgg	tactcatttg	catggtgtct	gcaaacagca	ccacctaaat	tgcacgtgt	1200
aatacacgta	gcagatttga	gtggaacata	atcaatatcc	gacactactt	gtttgccatg	1260
agactcacia	ggactatcag	aatagtaaaa	gaaaggcaat	tgctttaaat	tagtaaatgc	1320
acttttatcg	aaagctggag	tgtggaatgc	atgcttattc	acatacaaac	taccaccatc	1380
acagcctggg	aagttcaagt	ttgacaagac	tcttgtgtca	aacctacaca	caattgcatt	1440
ggctgggtaa	cgatcaacgt	tacaattcca	aaacaaacaa	acaccatcag	tgaatttatc	1500
gtgatgtgta	gcataagaat	agaagagttc	ctctattttg	taagctttgt	cactacatgg	1560
ctgagcatcg	tagaacttcc	attctacttc	agcctgaggc	acacacttga	tagcctttgg	1620
atttccaatg	tcatgaagaa	ctggaaactt	atcagcaagc	aatgcagact	tcacaaccat	1680
gtgttgtagt	tttctgcaag	cagaattaac	cctcagttca	tctcctataa	taggggtattc	1740
aacagaccaa	tcaacgcgct	taacaaagca	ctcatggact	gctaaacatc	tagtcatgat	1800
agcatcacaa	ctagccacat	gtgcatttcc	atgtacctgg	caatgttggt	catggttact	1860
ctgaagggtta	cccgtaaagc	cccactgctg	aacatcaatc	ataaatgggt	tatagacata	1920
gtcaaaaacc	acagaatgat	tccagcaggc	ataagtatct	gatgaagtag	aaaagcaagt	1980
tgcacgtttg	tcacacagac	aacacgttct	ttcaggtcca	atcttgacaa	agtacttcat	2040
tgatgtaagc	tcaaagccat	gcgccccaaag	gacga			2075

<210> 54
 <211> 1891
 <212> ADN
 <213> CORONAVIRUS

<400> 54

aagattcacc	acttaaatta	gaggatttta	tccctatgga	cagcacagt	aaaaattact	60
tcataacaga	tgcgcaaaca	ggttcatcaa	aatgtgtgtg	ttctgtgatt	gatcttttac	120
ttgatgactt	tgtcgagata	ataaagtcac	aagatttgtc	agtgatttca	aaagtgggtca	180
aggttacaat	tgactatgct	gaaatttcat	tcatgctttg	gtgtaaggat	ggacatgttg	240
aaaccttcta	cccaaaacta	caagcaagtc	aagcgtggca	accaggtgtt	gcgatgccta	300
acttgtagaa	gatgcaaaga	atgcttcttg	aaaagtgtga	ccttcagaat	tatggtgaaa	360
atgctgttat	accaaagga	ataatgatga	atgtcgcaaa	gtatactcaa	ctgtgtcaat	420
acttaaatac	acttacttta	gctgtaccct	acaacatgag	agttattcac	tttgggtgctg	480
gctctgataa	aggagtgtga	ccaggtacag	ctgtgctcag	acaatgggtg	ccaactggca	540
cactacttgt	cgattcagat	cttaatgact	tcgtctccga	cgcagattct	actttaattg	600
gagactgtgc	aacagtacat	acggctaata	aatgggacct	tattattagc	gatatgtatg	660
accctaggac	caaacatgtg	acaaaagaga	atgactctaa	agaaggggtt	ttcacttatc	720
tgtgtggatt	tataaagcaa	aaactagccc	tgggtgggtc	tatagctgta	aagataacag	780
agcattcttg	gaatgctgac	ctttacaagc	ttatgggccca	tttctcatgg	tggacagctt	840
ttgttacaaa	tgtaaatagc	tcatcatcgg	aagcattttt	aattggggct	aactatcttg	900
gcaagccgaa	ggaacaaatt	gatggctata	ccatgcatgc	taactacatt	ttctggagga	960
acacaaatcc	tatccagttg	tcttcctatt	cactctttga	catgagcaaa	tttctcttta	1020
aattaagagg	aactgctgta	atgtctctta	aggagaatca	aatcaatgat	atgattttatt	1080
ctcttctgga	aaaaggtagg	cttatcatta	gagaaaacaa	cagagttgtg	gtttcaagtg	1140
atattcttgt	taacaactaa	acgaacatgt	ttattttctt	attattttctt	actctcacta	1200
gtggtagtga	ccttgaccgg	tgcaccactt	ttgatgatgt	tcaagctcct	aattacactc	1260
aacatacttc	atctatgagg	ggggtttact	atcctgatga	aatttttaga	tcagacactc	1320
tttatttaac	tcaggattta	tttcttccat	tttattctaa	tgttacaggg	tttcatacta	1380
ttaatcatac	gtttggcaac	cctgtcatac	cttttaagga	tggtatttat	tttgctgcca	1440
cagagaaatc	aaatgttgtc	cgtggttggg	tttttggttc	taccatgaac	aacaagtcac	1500
agtcggtgat	tattattaac	aattctacta	atgttggtat	acgagcatgt	aactttgaat	1560
tgtgtgacaa	ccctttcttt	gctgtttcta	aacccatggg	tacacagaca	catactatga	1620
tattcgataa	tgcatttaat	tgcactttcg	agtacataat	tgatgccttt	tcgcttgatg	1680
tttcagaaaa	gtcaggtaat	tttaaacact	tacgagagtt	tgtgtttaaa	aataaagatg	1740
ggtttctcta	tgtttataag	ggctatcaac	ctatagatgt	agttcgtgat	ctaccttctg	1800
gttttaacac	tttgaaacct	atttttaagt	tgcctcttgg	tattaacatt	acaaatttta	1860
gagccattct	tacagccttt	tcacctgctc	a			1891

<210> 55
 <211> 32
 <212> ADN

<213> SECUENCIA ARTIFICIAL

<220>

<223> cebador N sentido

<400> 55

cccatatgtc tgataatgga cccaatcaa ac 32

<210> 56

<211> 32

<212> ADN

<213> secuencia artificial

<220>

<223> cebador N antisentido

<400> 56

cccccggtg cctgagtgga atcagcagaa gc 32

<210> 57

<211> 31

<212> ADN

<213> secuencia artificial

<220>

<223> cebador Sc sentido

<400> 57

cccatatgag tgacctgac cggcgacca c 31

<210> 58

<211> 30

<212> ADN

<213> secuencia artificial

<220>

<223> cebador SL sentido

<400> 58

cccatatgaa acctgcacc ccactgctc 30

<210> 59

<211> 33

<212> ADN

<213> cebador Sc et SL antisentido

<400> 59

cccccggtt taatatattg ctcatatattt ccc 33

<210> 60

<211> 16

<212> ADN

<213> cebador sentido serie 1

<400> 60

ggcatcgat ggggtg 16

<210> 61

<211> 16

<212> ADN

<213> cebador antisentido serie 2 (28774-28759)

<400>61

cagtttcacc acctcc 16

<210> 62

<211> 16
 <212> ADN
 <213> cebador sentido serie 2 (28375-28390)

<400> 62
 ggctactacc gaagag 16

<210> 63
 <211> 16
 <212> ADN
 <213> cebador antisentido serie 2 (28702-28687)

<400> 63
 aattaccgcg actacg 16

<210> 64
 <211> 26
 <212> ADN
 <213> sonda 1/serie 1 (28561-28586)

<400> 64
 ggcacccgca atcctaataa caatgc 26

<210> 65
 <211> 21
 <212> ADN
 <213> sonda 2/serie 1 (28588-28608)

<400> 65
 gccaccgtgc tacaacttc t 21

<210> 66
 <211> 23
 <212> ADN
 <213> sonda 1/serie 2 /sonda N/FL (28541-28563)

<400> 66
 atacacccaa agaccacatt ggc 23

<210> 67
 <211> 25
 <212> ADN
 <213> sonda 2/serie 2/sonda SRAS/N/LC705 (28565-28589)

<400> 67
 cccgcaatcc taataacaat gctgc 25

<210> 68
 <211> 30
 <212> ADN
 <213> secuencia artificial

<220>
 <223> cebador anclaje 14T

<400> 68
 agatgaattc ggtacctttt tttttttt 30

<210> 69
 <211> 13
 <212> PRT
 <213> secuencia artificial

<220>
 <223> péptido M2-14

ES 2 396 127 T3

<400> 69

Ala Asp Asn Gly Thr Ile Thr Val Glu Glu Leu Lys Gln
1 5 10

<210> 70

<211> 12

<212> PRT

<213> secuencia artificial

<220>

<223> péptido E1-12

<400> 70

Met Tyr Ser Phe Val Ser Glu Glu Thr Gly Thr Leu
1 5 10

<210> 71

<211> 24

<212> PRT

<213> secuencia artificial

<220>

<223> péptido E53-72

<400> 71

Lys Pro Thr Val Tyr Val Tyr Ser Arg Val Lys Asn Leu Asn Ser Ser

1 5 10 15

Glu Gly Val Pro Asp Leu Leu Val
20

<210> 72

<211> 153

<212> ADN

<213> CORONAVIRUS

<400> 72

gatattaggt ttttacctac ccaggaaaag ccaaccaacc tcgatctctt gtagatctgt	60
tctctaaacg aacttttaaaa tctgtgtagc tgtcgctcgg ctgcatgcct agtgcaccta	120
cgcagtataa acaataataa attttactgt cgt	153

<210> 73

<211> 410

<212> ADN

<213> CORONAVIRUS

<400> 73

ES 2 396 127 T3

```

ttctccagac aacttcaaaa ttccatgagt ggagcttctg ctgattcaac tcaggcataa      60
acactcatga tgaccacaca aggcagatgg gctatgtaaa cgttttcgca attccgttta      120
cgatacatag tctactcttg tgcagaatga attctcgtaa ctaaacagca caagtagggt      180
tagttaactt taatctcaca tagcaatctt taatcaatgt gtaacattag ggaggacttg      240
aaagagccac cacattttca tcgaggccac gcggagtacg atcgagggta cagtgaataa      300
tgctagggag agctgcctat atggaagagc cctaattgtgt aaaattaatt ttagtagtgc      360
tatcccatg  tgattttaat agcttcttag gagaatgaca aaaaaaaaaa      410

```

<210> 74

<211> 4382

<212> PRT

<213> CORONAVIRUS

<400> 74

```

Met Glu Ser Leu Val Leu Gly Val Asn Glu Lys Thr His Val Gln Leu
1           5           10          15

```

```

Ser Leu Pro Val Leu Gln Val Arg Asp Val Leu Val Arg Gly Phe Gly

```

ES 2 396 127 T3

20										25					30															
Asp	Ser	Val	Glu	Glu	Ala	Leu	Ser	Glu	Ala	Arg	Glu	His	Leu	Lys	Asn															
		35					40					45																		
Gly	Thr	Cys	Gly	Leu	Val	Glu	Leu	Glu	Lys	Gly	Val	Leu	Pro	Gln	Leu															
	50					55					60																			
Glu	Gln	Pro	Tyr	Val	Phe	Ile	Lys	Arg	Ser	Asp	Ala	Leu	Ser	Thr	Asn															
65					70					75					80															
His	Gly	His	Lys	Val	Val	Glu	Leu	Val	Ala	Glu	Met	Asp	Gly	Ile	Gln															
				85					90					95																
Tyr	Gly	Arg	Ser	Gly	Ile	Thr	Leu	Gly	Val	Leu	Val	Pro	His	Val	Gly															
			100					105					110																	
Glu	Thr	Pro	Ile	Ala	Tyr	Arg	Asn	Val	Leu	Leu	Arg	Lys	Asn	Gly	Asn															
		115					120					125																		
Lys	Gly	Ala	Gly	Gly	His	Ser	Tyr	Gly	Ile	Asp	Leu	Lys	Ser	Tyr	Asp															
	130					135					140																			
Leu	Gly	Asp	Glu	Leu	Gly	Thr	Asp	Pro	Ile	Glu	Asp	Tyr	Glu	Gln	Asn															
145					150					155					160															
Trp	Asn	Thr	Lys	His	Gly	Ser	Gly	Ala	Leu	Arg	Glu	Leu	Thr	Arg	Glu															
				165					170					175																
Leu	Asn	Gly	Gly	Ala	Val	Thr	Arg	Tyr	Val	Asp	Asn	Asn	Phe	Cys	Gly															
			180					185					190																	
Pro	Asp	Gly	Tyr	Pro	Leu	Asp	Cys	Ile	Lys	Asp	Phe	Leu	Ala	Arg	Ala															
		195					200					205																		
Gly	Lys	Ser	Met	Cys	Thr	Leu	Ser	Glu	Gln	Leu	Asp	Tyr	Ile	Glu	Ser															
	210					215					220																			
Lys	Arg	Gly	Val	Tyr	Cys	Cys	Arg	Asp	His	Glu	His	Glu	Ile	Ala	Trp															
225					230					235					240															
Phe	Thr	Glu	Arg	Ser	Asp	Lys	Ser	Tyr	Glu	His	Gln	Thr	Pro	Phe	Glu															
				245					250					255																
Ile	Lys	Ser	Ala	Lys	Lys	Phe	Asp	Thr	Phe	Lys	Gly	Glu	Cys	Pro	Lys															
			260					265					270																	
Phe	Val	Phe	Pro	Leu	Asn	Ser	Lys	Val	Lys	Val	Ile	Gln	Pro	Arg	Val															
		275					280					285																		
Glu	Lys	Lys	Lys	Thr	Glu	Gly	Phe	Met	Gly	Arg	Ile	Arg	Ser	Val	Tyr															

ES 2 396 127 T3

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

565										570					575					
Ser	Leu	Arg	Leu	Val	Asp	Ala	Met	Val	Tyr	Thr	Ser	Asp	Leu	Leu	Thr					
			580					585					590							
Asn	Ser	Val	Ile	Ile	Met	Ala	Tyr	Val	Thr	Gly	Gly	Leu	Val	Gln	Gln					
		595					600					605								
Thr	Ser	Gln	Trp	Leu	Ser	Asn	Leu	Leu	Gly	Thr	Thr	Val	Glu	Lys	Leu					
	610					615					620									
Arg	Pro	Ile	Phe	Glu	Trp	Ile	Glu	Ala	Lys	Leu	Ser	Ala	Gly	Val	Glu					
625					630					635					640					
Phe	Leu	Lys	Asp	Ala	Trp	Glu	Ile	Leu	Lys	Phe	Leu	Ile	Thr	Gly	Val					
				645					650					655						
Phe	Asp	Ile	Val	Lys	Gly	Gln	Ile	Gln	Val	Ala	Ser	Asp	Asn	Ile	Lys					
			660					665					670							
Asp	Cys	Val	Lys	Cys	Phe	Ile	Asp	Val	Val	Asn	Lys	Ala	Leu	Glu	Met					
		675					680					685								
Cys	Ile	Asp	Gln	Val	Thr	Ile	Ala	Gly	Ala	Lys	Leu	Arg	Ser	Leu	Asn					
	690					695					700									
Leu	Gly	Glu	Val	Phe	Ile	Ala	Gln	Ser	Lys	Gly	Leu	Tyr	Arg	Gln	Cys					
705					710					715					720					
Ile	Arg	Gly	Lys	Glu	Gln	Leu	Gln	Leu	Leu	Met	Pro	Leu	Lys	Ala	Pro					
				725					730					735						
Lys	Glu	Val	Thr	Phe	Leu	Glu	Gly	Asp	Ser	His	Asp	Thr	Val	Leu	Thr					
			740					745					750							
Ser	Glu	Glu	Val	Val	Leu	Lys	Asn	Gly	Glu	Leu	Glu	Ala	Leu	Glu	Thr					
		755					760					765								
Pro	Val	Asp	Ser	Phe	Thr	Asn	Gly	Ala	Ile	Val	Gly	Thr	Pro	Val	Cys					
	770					775					780									
Val	Asn	Gly	Leu	Met	Leu	Leu	Glu	Ile	Lys	Asp	Lys	Glu	Gln	Tyr	Cys					
785					790					795					800					
Ala	Leu	Ser	Pro	Gly	Leu	Leu	Ala	Thr	Asn	Asn	Val	Phe	Arg	Leu	Lys					
				805					810					815						
Gly	Gly	Ala	Pro	Ile	Lys	Gly	Val	Thr	Phe	Gly	Glu	Asp	Thr	Val	Trp					
			820					825					830							
Glu	Val	Gln	Gly	Tyr	Lys	Asn	Val	Arg	Ile	Thr	Phe	Glu	Leu	Asp	Glu					

835	840	845
Arg Val Asp Lys Val Leu Asn Glu Lys Cys Ser Val Tyr Thr Val Glu 850 855 860		
Ser Gly Thr Glu Val Thr Glu Phe Ala Cys Val Val Ala Glu Ala Val 865 870 875 880		
Val Lys Thr Leu Gln Pro Val Ser Asp Leu Leu Thr Asn Met Gly Ile 885 890 895		
Asp Leu Asp Glu Trp Ser Val Ala Thr Phe Tyr Leu Phe Asp Asp Ala 900 905 910		
Gly Glu Glu Asn Phe Ser Ser Arg Met Tyr Cys Ser Phe Tyr Pro Pro 915 920 925		
Asp Glu Glu Glu Glu Asp Asp Ala Glu Cys Glu Glu Glu Glu Ile Asp 930 935 940		
Glu Thr Cys Glu His Glu Tyr Gly Thr Glu Asp Asp Tyr Gln Gly Leu 945 950 955 960		
Pro Leu Glu Phe Gly Ala Ser Ala Glu Thr Val Arg Val Glu Glu Glu 965 970 975		
Glu Glu Glu Asp Trp Leu Asp Asp Thr Thr Glu Gln Ser Glu Ile Glu 980 985 990		
Pro Glu Pro Glu Pro Thr Pro Glu Glu Pro Val Asn Gln Phe Thr Gly 995 1000 1005		
Tyr Leu Lys Leu Thr Asp Asn Val Ala Ile Lys Cys Val Asp Ile 1010 1015 1020		
Val Lys Glu Ala Gln Ser Ala Asn Pro Met Val Ile Val Asn Ala 1025 1030 1035		
Ala Asn Ile His Leu Lys His Gly Gly Gly Val Ala Gly Ala Leu 1040 1045 1050		
Asn Lys Ala Thr Asn Gly Ala Met Gln Lys Glu Ser Asp Asp Tyr 1055 1060 1065		
Ile Lys Leu Asn Gly Pro Leu Thr Val Gly Gly Ser Cys Leu Leu 1070 1075 1080		
Ser Gly His Asn Leu Ala Lys Lys Cys Leu His Val Val Gly Pro 1085 1090 1095		
Asn Leu Asn Ala Gly Glu Asp Ile Gln Leu Leu Lys Ala Ala Tyr		

1100	1105	1110
Glu Asn Phe Asn Ser Gln	Asp Ile Leu Leu Ala	Pro Leu Leu Ser
1115	1120	1125
Ala Gly Ile Phe Gly Ala	Lys Pro Leu Gln Ser	Leu Gln Val Cys
1130	1135	1140
Val Gln Thr Val Arg Thr	Gln Val Tyr Ile Ala	Val Asn Asp Lys
1145	1150	1155
Ala Leu Tyr Glu Gln Val	Val Met Asp Tyr Leu	Asp Asn Leu Lys
1160	1165	1170
Pro Arg Val Glu Ala Pro	Lys Gln Glu Glu Pro	Pro Asn Thr Glu
1175	1180	1185
Asp Ser Lys Thr Glu Glu	Lys Ser Val Val Gln	Lys Pro Val Asp
1190	1195	1200
Val Lys Pro Lys Ile Lys	Ala Cys Ile Asp Glu	Val Thr Thr Thr
1205	1210	1215
Leu Glu Glu Thr Lys Phe	Leu Thr Asn Lys Leu	Leu Leu Phe Ala
1220	1225	1230
Asp Ile Asn Gly Lys Leu	Tyr His Asp Ser Gln	Asn Met Leu Arg
1235	1240	1245
Gly Glu Asp Met Ser Phe	Leu Glu Lys Asp Ala	Pro Tyr Met Val
1250	1255	1260
Gly Asp Val Ile Thr Ser	Gly Asp Ile Thr Cys	Val Val Ile Pro
1265	1270	1275
Ser Lys Lys Ala Gly Gly	Thr Thr Glu Met Leu	Ser Arg Ala Leu
1280	1285	1290
Lys Lys Val Pro Val Asp	Glu Tyr Ile Thr Thr	Tyr Pro Gly Gln
1295	1300	1305
Gly Cys Ala Gly Tyr Thr	Leu Glu Glu Ala Lys	Thr Ala Leu Lys
1310	1315	1320
Lys Cys Lys Ser Ala Phe	Tyr Val Leu Pro Ser	Glu Ala Pro Asn
1325	1330	1335
Ala Lys Glu Glu Ile Leu	Gly Thr Val Ser Trp	Asn Leu Arg Glu
1340	1345	1350
Met Leu Ala His Ala Glu	Glu Thr Arg Lys Leu	Met Pro Ile Cys

1355		1360		1365
Met Asp Val Arg Ala Ile	Met Ala Thr Ile Gln	Arg Lys Tyr Lys		
1370	1375	1380		
Gly Ile Lys Ile Gln Glu	Gly Ile Val Asp Tyr	Gly Val Arg Phe		
1385	1390	1395		
Phe Phe Tyr Thr Ser Lys	Glu Pro Val Ala Ser	Ile Ile Thr Lys		
1400	1405	1410		
Leu Asn Ser Leu Asn Glu	Pro Leu Val Thr Met	Pro Ile Gly Tyr		
1415	1420	1425		
Val Thr His Gly Phe Asn	Leu Glu Glu Ala Ala	Arg Cys Met Arg		
1430	1435	1440		
Ser Leu Lys Ala Pro Ala	Val Val Ser Val Ser	Ser Pro Asp Ala		
1445	1450	1455		
Val Thr Thr Tyr Asn Gly	Tyr Leu Thr Ser Ser	Ser Lys Thr Ser		
1460	1465	1470		
Glu Glu His Phe Val Glu	Thr Val Ser Leu Ala	Gly Ser Tyr Arg		
1475	1480	1485		
Asp Trp Ser Tyr Ser Gly	Gln Arg Thr Glu Leu	Gly Val Glu Phe		
1490	1495	1500		
Leu Lys Arg Gly Asp Lys	Ile Val Tyr His Thr	Leu Glu Ser Pro		
1505	1510	1515		
Val Glu Phe His Leu Asp	Gly Glu Val Leu Ser	Leu Asp Lys Leu		
1520	1525	1530		
Lys Ser Leu Leu Ser Leu	Arg Glu Val Lys Thr	Ile Lys Val Phe		
1535	1540	1545		
Thr Thr Val Asp Asn Thr	Asn Leu His Thr Gln	Leu Val Asp Met		
1550	1555	1560		
Ser Met Thr Tyr Gly Gln	Gln Phe Gly Pro Thr	Tyr Leu Asp Gly		
1565	1570	1575		
Ala Asp Val Thr Lys Ile	Lys Pro His Val Asn	His Glu Gly Lys		
1580	1585	1590		
Thr Phe Phe Val Leu Pro	Ser Asp Asp Thr Leu	Arg Ser Glu Ala		
1595	1600	1605		
Phe Glu Tyr Tyr His Thr	Leu Asp Glu Ser Phe	Leu Gly Arg Tyr		

1610						1615					1620			
Met	Ser	Ala	Leu	Asn	His	Thr	Lys	Lys	Trp	Lys	Phe	Pro	Gln	Val
	1625					1630					1635			
Gly	Gly	Leu	Thr	Ser	Ile	Lys	Trp	Ala	Asp	Asn	Asn	Cys	Tyr	Leu
	1640					1645					1650			
Ser	Ser	Val	Leu	Leu	Ala	Leu	Gln	Gln	Leu	Glu	Val	Lys	Phe	Asn
	1655					1660					1665			
Ala	Pro	Ala	Leu	Gln	Glu	Ala	Tyr	Tyr	Arg	Ala	Arg	Ala	Gly	Asp
	1670					1675					1680			
Ala	Ala	Asn	Phe	Cys	Ala	Leu	Ile	Leu	Ala	Tyr	Ser	Asn	Lys	Thr
	1685					1690					1695			
Val	Gly	Glu	Leu	Gly	Asp	Val	Arg	Glu	Thr	Met	Thr	His	Leu	Leu
	1700					1705					1710			
Gln	His	Ala	Asn	Leu	Glu	Ser	Ala	Lys	Arg	Val	Leu	Asn	Val	Val
	1715					1720					1725			
Cys	Lys	His	Cys	Gly	Gln	Lys	Thr	Thr	Thr	Leu	Thr	Gly	Val	Glu
	1730					1735					1740			
Ala	Val	Met	Tyr	Met	Gly	Thr	Leu	Ser	Tyr	Asp	Asn	Leu	Lys	Thr
	1745					1750					1755			
Gly	Val	Ser	Ile	Pro	Cys	Val	Cys	Gly	Arg	Asp	Ala	Thr	Gln	Tyr
	1760					1765					1770			
Leu	Val	Gln	Gln	Glu	Ser	Ser	Phe	Val	Met	Met	Ser	Ala	Pro	Pro
	1775					1780					1785			
Ala	Glu	Tyr	Lys	Leu	Gln	Gln	Gly	Thr	Phe	Leu	Cys	Ala	Asn	Glu
	1790					1795					1800			
Tyr	Thr	Gly	Asn	Tyr	Gln	Cys	Gly	His	Tyr	Thr	His	Ile	Thr	Ala
	1805					1810					1815			
Lys	Glu	Thr	Leu	Tyr	Arg	Ile	Asp	Gly	Ala	His	Leu	Thr	Lys	Met
	1820					1825					1830			
Ser	Glu	Tyr	Lys	Gly	Pro	Val	Thr	Asp	Val	Phe	Tyr	Lys	Glu	Thr
	1835					1840					1845			
Ser	Tyr	Thr	Thr	Thr	Ile	Lys	Pro	Val	Ser	Tyr	Lys	Leu	Asp	Gly
	1850					1855					1860			
Val	Thr	Tyr	Thr	Glu	Ile	Glu	Pro	Lys	Leu	Asp	Gly	Tyr	Tyr	Lys

ES 2 396 127 T3

1865		1870		1875
Lys Asp 1880	Asn Ala Tyr Tyr	Thr 1885	Glu Gln Pro Ile	Asp 1890 Leu Val Pro
Thr Gln 1895	Pro Leu Pro Asn	Ala 1900	Ser Phe Asp Asn	Phe 1905 Lys Leu Thr
Cys Ser 1910	Asn Thr Lys Phe	Ala 1915	Asp Asp Leu Asn	Gln 1920 Met Thr Gly
Phe Thr 1925	Lys Pro Ala Ser	Arg 1930	Glu Leu Ser Val	Thr 1935 Phe Phe Pro
Asp Leu 1940	Asn Gly Asp Val	Val 1945	Ala Ile Asp Tyr	Arg 1950 His Tyr Ser
Ala Ser 1955	Phe Lys Lys Gly	Ala 1960	Lys Leu Leu His	Lys 1965 Pro Ile Val
Trp His 1970	Ile Asn Gln Ala	Thr 1975	Thr Lys Thr Thr	Phe 1980 Lys Pro Asn
Thr Trp 1985	Cys Leu Arg Cys	Leu 1990	Trp Ser Thr Lys	Pro 1995 Val Asp Thr
Ser Asn 2000	Ser Phe Glu Val	Leu 2005	Ala Val Glu Asp	Thr 2010 Gln Gly Met
Asp Asn 2015	Leu Ala Cys Glu	Ser 2020	Gln Gln Pro Thr	Ser 2025 Glu Glu Val
Val Glu 2030	Asn Pro Thr Ile	Gln 2035	Lys Glu Val Ile	Glu 2040 Cys Asp Val
Lys Thr 2045	Thr Glu Val Val	Gly 2050	Asn Val Ile Leu	Lys 2055 Pro Ser Asp
Glu Gly 2060	Val Lys Val Thr	Gln 2065	Glu Leu Gly His	Glu 2070 Asp Leu Met
Ala Ala 2075	Tyr Val Glu Asn	Thr 2080	Ser Ile Thr Ile	Lys 2085 Lys Pro Asn
Glu Leu 2090	Ser Leu Ala Leu	Gly 2095	Leu Lys Thr Ile	Ala 2100 Thr His Gly
Ile Ala 2105	Ala Ile Asn Ser	Val 2110	Pro Trp Ser Lys	Ile 2115 Leu Ala Tyr
Val Lys	Pro Phe Leu Gly	Gln	Ala Ala Ile Thr	Thr Ser Asn Cys

ES 2 396 127 T3

2120	2125	2130
Ala Lys Arg Leu Ala Gln	Arg Val Phe Asn Asn	Tyr Met Pro Tyr
2135	2140	2145
Val Phe Thr Leu Leu Phe	Gln Leu Cys Thr Phe	Thr Lys Ser Thr
2150	2155	2160
Asn Ser Arg Ile Arg Ala	Ser Leu Pro Thr Thr	Ile Ala Lys Asn
2165	2170	2175
Ser Val Lys Ser Val Ala	Lys Leu Cys Leu Asp	Ala Gly Ile Asn
2180	2185	2190
Tyr Val Lys Ser Pro Lys	Phe Ser Lys Leu Phe	Thr Ile Ala Met
2195	2200	2205
Trp Leu Leu Leu Ser	Ile Cys Leu Gly Ser	Leu Ile Cys Val
2210	2215	2220
Thr Ala Ala Phe Gly Val	Leu Leu Ser Asn Phe	Gly Ala Pro Ser
2225	2230	2235
Tyr Cys Asn Gly Val Arg	Glu Leu Tyr Leu Asn	Ser Ser Asn Val
2240	2245	2250
Thr Thr Met Asp Phe Cys	Glu Gly Ser Phe Pro	Cys Ser Ile Cys
2255	2260	2265
Leu Ser Gly Leu Asp Ser	Leu Asp Ser Tyr Pro	Ala Leu Glu Thr
2270	2275	2280
Ile Gln Val Thr Ile Ser	Ser Tyr Lys Leu Asp	Leu Thr Ile Leu
2285	2290	2295
Gly Leu Ala Ala Glu Trp	Val Leu Ala Tyr Met	Leu Phe Thr Lys
2300	2305	2310
Phe Phe Tyr Leu Leu Gly	Leu Ser Ala Ile Met	Gln Val Phe Phe
2315	2320	2325
Gly Tyr Phe Ala Ser His	Phe Ile Ser Asn Ser	Trp Leu Met Trp
2330	2335	2340
Phe Ile Ile Ser Ile Val	Gln Met Ala Pro Val	Ser Ala Met Val
2345	2350	2355
Arg Met Tyr Ile Phe Phe	Ala Ser Phe Tyr Tyr	Ile Trp Lys Ser
2360	2365	2370
Tyr Val His Ile Met Asp	Gly Cys Thr Ser Ser	Thr Cys Met Met

ES 2 396 127 T3

2375						2380						2385
Cys Tyr	Lys Arg Asn Arg	Ala	Thr Arg Val Glu	Cys	Thr Thr Ile							
2390		2395		2400								
Val Asn	Gly Met Lys Arg	Ser	Phe Tyr Val Tyr	Ala	Asn Gly Gly							
2405		2410		2415								
Arg Gly	Phe Cys Lys Thr	His	Asn Trp Asn Cys	Leu	Asn Cys Asp							
2420		2425		2430								
Thr Phe	Cys Thr Gly Ser	Thr	Phe Ile Ser Asp	Glu	Val Ala Arg							
2435		2440		2445								
Asp Leu	Ser Leu Gln Phe	Lys	Arg Pro Ile Asn	Pro	Thr Asp Gln							
2450		2455		2460								
Ser Ser	Tyr Ile Val Asp	Ser	Val Ala Val Lys	Asn	Gly Ala Leu							
2465		2470		2475								
His Leu	Tyr Phe Asp Lys	Ala	Gly Gln Lys Thr	Tyr	Glu Arg His							
2480		2485		2490								
Pro Leu	Ser His Phe Val	Asn	Leu Asp Asn Leu	Arg	Ala Asn Asn							
2495		2500		2505								
Thr Lys	Gly Ser Leu Pro	Ile	Asn Val Ile Val	Phe	Asp Gly Lys							
2510		2515		2520								
Ser Lys	Cys Asp Glu Ser	Ala	Ser Lys Ser Ala	Ser	Val Tyr Tyr							
2525		2530		2535								
Ser Gln	Leu Met Cys Gln	Pro	Ile Leu Leu Leu	Asp	Gln Ala Leu							
2540		2545		2550								
Val Ser	Asp Val Gly Asp	Ser	Thr Glu Val Ser	Val	Lys Met Phe							
2555		2560		2565								
Asp Ala	Tyr Val Asp Thr	Phe	Ser Ala Thr Phe	Ser	Val Pro Met							
2570		2575		2580								
Glu Lys	Leu Lys Ala Leu	Val	Ala Thr Ala His	Ser	Glu Leu Ala							
2585		2590		2595								
Lys Gly	Val Ala Leu Asp	Gly	Val Leu Ser Thr	Phe	Val Ser Ala							
2600		2605		2610								
Ala Arg	Gln Gly Val Val	Asp	Thr Asp Val Asp	Thr	Lys Asp Val							
2615		2620		2625								
Ile Glu	Cys Leu Lys Leu	Ser	His His Ser Asp	Leu	Glu Val Thr							

2630						2635						2640		
Gly	Asp	Ser	Cys	Asn	Asn	Phe	Met	Leu	Thr	Tyr	Asn	Lys	Val	Glu
	2645					2650					2655			
Asn	Met	Thr	Pro	Arg	Asp	Leu	Gly	Ala	Cys	Ile	Asp	Cys	Asn	Ala
	2660					2665					2670			
Arg	His	Ile	Asn	Ala	Gln	Val	Ala	Lys	Ser	His	Asn	Val	Ser	Leu
	2675					2680					2685			
Ile	Trp	Asn	Val	Lys	Asp	Tyr	Met	Ser	Leu	Ser	Glu	Gln	Leu	Arg
	2690					2695					2700			
Lys	Gln	Ile	Arg	Ser	Ala	Ala	Lys	Lys	Asn	Asn	Ile	Pro	Phe	Arg
	2705					2710					2715			
Leu	Thr	Cys	Ala	Thr	Thr	Arg	Gln	Val	Val	Asn	Val	Ile	Thr	Thr
	2720					2725					2730			
Lys	Ile	Ser	Leu	Lys	Gly	Gly	Lys	Ile	Val	Ser	Thr	Cys	Phe	Lys
	2735					2740					2745			
Leu	Met	Leu	Lys	Ala	Thr	Leu	Leu	Cys	Val	Leu	Ala	Ala	Leu	Val
	2750					2755					2760			
Cys	Tyr	Ile	Val	Met	Pro	Val	His	Thr	Leu	Ser	Ile	His	Asp	Gly
	2765					2770					2775			
Tyr	Thr	Asn	Glu	Ile	Ile	Gly	Tyr	Lys	Ala	Ile	Gln	Asp	Gly	Val
	2780					2785					2790			
Thr	Arg	Asp	Ile	Ile	Ser	Thr	Asp	Asp	Cys	Phe	Ala	Asn	Lys	His
	2795					2800					2805			
Ala	Gly	Phe	Asp	Ala	Trp	Phe	Ser	Gln	Arg	Gly	Gly	Ser	Tyr	Lys
	2810					2815					2820			
Asn	Asp	Lys	Ser	Cys	Pro	Val	Val	Ala	Ala	Ile	Ile	Thr	Arg	Glu
	2825					2830					2835			
Ile	Gly	Phe	Ile	Val	Pro	Gly	Leu	Pro	Gly	Thr	Val	Leu	Arg	Ala
	2840					2845					2850			
Ile	Asn	Gly	Asp	Phe	Leu	His	Phe	Leu	Pro	Arg	Val	Phe	Ser	Ala
	2855					2860					2865			
Val	Gly	Asn	Ile	Cys	Tyr	Thr	Pro	Ser	Lys	Leu	Ile	Glu	Tyr	Ser
	2870					2875					2880			
Asp	Phe	Ala	Thr	Ser	Ala	Cys	Val	Leu	Ala	Ala	Glu	Cys	Thr	Ile

ES 2 396 127 T3

2885						2890						2895
Phe	Lys	Asp	Ala	Met	Gly	Lys	Pro	Val	Pro	Tyr	Cys	Tyr
2900						2905					2910	Asp
												Thr
Asn	Leu	Leu	Glu	Gly	Ser	Ile	Ser	Tyr	Ser	Glu	Leu	Arg
2915						2920					2925	Pro
												Asp
Thr	Arg	Tyr	Val	Leu	Met	Asp	Gly	Ser	Ile	Ile	Gln	Phe
2930						2935					2940	Pro
												Asn
Thr	Tyr	Leu	Glu	Gly	Ser	Val	Arg	Val	Val	Thr	Thr	Phe
2945						2950					2955	Asp
												Ala
Glu	Tyr	Cys	Arg	His	Gly	Thr	Cys	Glu	Arg	Ser	Glu	Val
2960						2965					2970	Gly
												Ile
Cys	Leu	Ser	Thr	Ser	Gly	Arg	Trp	Val	Leu	Asn	Asn	Glu
2975						2980					2985	His
												Tyr
Arg	Ala	Leu	Ser	Gly	Val	Phe	Cys	Gly	Val	Asp	Ala	Met
2990						2995					3000	Asn
												Leu
Ile	Ala	Asn	Ile	Phe	Thr	Pro	Leu	Val	Gln	Pro	Val	Gly
3005						3010					3015	Ala
												Leu
Asp	Val	Ser	Ala	Ser	Val	Val	Ala	Gly	Gly	Ile	Ile	Ala
3020						3025					3030	Ile
												Leu
Val	Thr	Cys	Ala	Ala	Tyr	Tyr	Phe	Met	Lys	Phe	Arg	Arg
3035						3040					3045	Val
												Phe
Gly	Glu	Tyr	Asn	His	Val	Val	Ala	Ala	Asn	Ala	Leu	Leu
3050						3055					3060	Phe
												Leu
Met	Ser	Phe	Thr	Ile	Leu	Cys	Leu	Val	Pro	Ala	Tyr	Ser
3065						3070					3075	Phe
												Leu
Pro	Gly	Val	Tyr	Ser	Val	Phe	Tyr	Leu	Tyr	Leu	Thr	Phe
3080						3085					3090	Tyr
												Phe
Thr	Asn	Asp	Val	Ser	Phe	Leu	Ala	His	Leu	Gln	Trp	Phe
3095						3100					3105	Ala
												Met
Phe	Ser	Pro	Ile	Val	Pro	Phe	Trp	Ile	Thr	Ala	Ile	Tyr
3110						3115					3120	Val
												Phe
Cys	Ile	Ser	Leu	Lys	His	Cys	His	Trp	Phe	Phe	Asn	Asn
3125						3130					3135	Tyr
												Leu
Arg	Lys	Arg	Val	Met	Phe	Asn	Gly	Val	Thr	Phe	Ser	Thr
												Phe
												Glu

3140		3145		3150
Glu Ala	Ala Leu Cys Thr	Phe Leu Leu Asn Lys	Glu Met Tyr Leu	
3155		3160		3165
Lys Leu	Arg Ser Glu Thr	Leu Leu Pro Leu Thr	Gln Tyr Asn Arg	
3170		3175		3180
Tyr Leu	Ala Leu Tyr Asn	Lys Tyr Lys Tyr Phe	Ser Gly Ala Leu	
3185		3190		3195
Asp Thr	Thr Ser Tyr Arg	Glu Ala Ala Cys Cys	His Leu Ala Lys	
3200		3205		3210
Ala Leu	Asn Asp Phe Ser	Asn Ser Gly Ala Asp	Val Leu Tyr Gln	
3215		3220		3225
Pro Pro	Gln Thr Ser Ile	Thr Ser Ala Val Leu	Gln Ser Gly Phe	
3230		3235		3240
Arg Lys	Met Ala Phe Pro	Ser Gly Lys Val Glu	Gly Cys Met Val	
3245		3250		3255
Gln Val	Thr Cys Gly Thr	Thr Thr Leu Asn Gly	Leu Trp Leu Asp	
3260		3265		3270
Asp Thr	Val Tyr Cys Pro	Arg His Val Ile Cys	Thr Ala Glu Asp	
3275		3280		3285
Met Leu	Asn Pro Asn Tyr	Glu Asp Leu Leu Ile	Arg Lys Ser Asn	
3290		3295		3300
His Ser	Phe Leu Val Gln	Ala Gly Asn Val Gln	Leu Arg Val Ile	
3305		3310		3315
Gly His	Ser Met Gln Asn	Cys Leu Leu Arg Leu	Lys Val Asp Thr	
3320		3325		3330
Ser Asn	Pro Lys Thr Pro	Lys Tyr Lys Phe Val	Arg Ile Gln Pro	
3335		3340		3345
Gly Gln	Thr Phe Ser Val	Leu Ala Cys Tyr Asn	Gly Ser Pro Ser	
3350		3355		3360
Gly Val	Tyr Gln Cys Ala	Met Arg Pro Asn His	Thr Ile Lys Gly	
3365		3370		3375
Ser Phe	Leu Asn Gly Ser	Cys Gly Ser Val Gly	Phe Asn Ile Asp	
3380		3385		3390
Tyr Asp	Cys Val Ser Phe	Cys Tyr Met His His	Met Glu Leu Pro	

ES 2 396 127 T3

3395						3400						3405
Thr	Gly	Val	His	Ala	Gly	Thr	Asp	Leu	Glu	Gly	Lys	Phe Tyr Gly
	3410						3415				3420	
Pro	Phe	Val	Asp	Arg	Gln	Thr	Ala	Gln	Ala	Ala	Gly	Thr Asp Thr
	3425						3430				3435	
Thr	Ile	Thr	Leu	Asn	Val	Leu	Ala	Trp	Leu	Tyr	Ala	Ala Val Ile
	3440						3445				3450	
Asn	Gly	Asp	Arg	Trp	Phe	Leu	Asn	Arg	Phe	Thr	Thr	Thr Leu Asn
	3455						3460				3465	
Asp	Phe	Asn	Leu	Val	Ala	Met	Lys	Tyr	Asn	Tyr	Glu	Pro Leu Thr
	3470						3475				3480	
Gln	Asp	His	Val	Asp	Ile	Leu	Gly	Pro	Leu	Ser	Ala	Gln Thr Gly
	3485						3490				3495	
Ile	Ala	Val	Leu	Asp	Met	Cys	Ala	Ala	Leu	Lys	Glu	Leu Leu Gln
	3500						3505				3510	
Asn	Gly	Met	Asn	Gly	Arg	Thr	Ile	Leu	Gly	Ser	Thr	Ile Leu Glu
	3515						3520				3525	
Asp	Glu	Phe	Thr	Pro	Phe	Asp	Val	Val	Arg	Gln	Cys	Ser Gly Val
	3530						3535				3540	
Thr	Phe	Gln	Gly	Lys	Phe	Lys	Lys	Ile	Val	Lys	Gly	Thr His His
	3545						3550				3555	
Trp	Met	Leu	Leu	Thr	Phe	Leu	Thr	Ser	Leu	Leu	Ile	Leu Val Gln
	3560						3565				3570	
Ser	Thr	Gln	Trp	Ser	Leu	Phe	Phe	Phe	Val	Tyr	Glu	Asn Ala Phe
	3575						3580				3585	
Leu	Pro	Phe	Thr	Leu	Gly	Ile	Met	Ala	Ile	Ala	Ala	Cys Ala Met
	3590						3595				3600	
Leu	Leu	Val	Lys	His	Lys	His	Ala	Phe	Leu	Cys	Leu	Phe Leu Leu
	3605						3610				3615	
Pro	Ser	Leu	Ala	Thr	Val	Ala	Tyr	Phe	Asn	Met	Val	Tyr Met Pro
	3620						3625				3630	
Ala	Ser	Trp	Val	Met	Arg	Ile	Met	Thr	Trp	Leu	Glu	Leu Ala Asp
	3635						3640				3645	
Thr	Ser	Leu	Ser	Gly	Tyr	Arg	Leu	Lys	Asp	Cys	Val	Met Tyr Ala

3650		3655		3660
Ser Ala 3665	Leu Val Leu Leu	Ile 3670	Leu Met Thr Ala	Arg 3675 Thr Val Tyr
Asp Asp 3680	Ala Ala Arg Arg	Val 3685	Trp Thr Leu Met	Asn 3690 Val Ile Thr
Leu Val 3695	Tyr Lys Val Tyr	Tyr 3700	Gly Asn Ala Leu	Asp 3705 Gln Ala Ile
Ser Met 3710	Trp Ala Leu Val	Ile 3715	Ser Val Thr Ser	Asn 3720 Tyr Ser Gly
Val Val 3725	Thr Thr Ile Met	Phe 3730	Leu Ala Arg Ala	Ile 3735 Val Phe Val
Cys Val 3740	Glu Tyr Tyr Pro	Leu 3745	Leu Phe Ile Thr	Gly 3750 Asn Thr Leu
Gln Cys 3755	Ile Met Leu Val	Tyr 3760	Cys Phe Leu Gly	Tyr 3765 Cys Cys Cys
Cys Tyr 3770	Phe Gly Leu Phe	Cys 3775	Leu Leu Asn Arg	Tyr 3780 Phe Arg Leu
Thr Leu 3785	Gly Val Tyr Asp	Tyr 3790	Leu Val Ser Thr	Gln 3795 Glu Phe Arg
Tyr Met 3800	Asn Ser Gln Gly	Leu 3805	Leu Pro Pro Lys	Ser 3810 Ser Ile Asp
Ala Phe 3815	Lys Leu Asn Ile	Lys 3820	Leu Leu Gly Ile	Gly 3825 Gly Lys Pro
Cys Ile 3830	Lys Val Ala Thr	Val 3835	Gln Ser Lys Met	Ser 3840 Asp Val Lys
Cys Thr 3845	Ser Val Val Leu	Leu 3850	Ser Val Leu Gln	Gln 3855 Leu Arg Val
Glu Ser 3860	Ser Ser Lys Leu	Trp 3865	Ala Gln Cys Val	Gln 3870 Leu His Asn
Asp Ile 3875	Leu Leu Ala Lys	Asp 3880	Thr Thr Glu Ala	Phe 3885 Glu Lys Met
Val Ser 3890	Leu Leu Ser Val	Leu 3895	Leu Ser Met Gln	Gly 3900 Ala Val Asp
Ile Asn	Arg Leu Cys Glu Glu	Met Leu Asp Asn Arg	Ala Thr Leu	

ES 2 396 127 T3

3905		3910		3915
Gln Ala 3920	Ile Ala Ser Glu	Phe 3925	Ser Ser Leu Pro	Ser 3930 Tyr Ala Ala
Tyr Ala 3935	Thr Ala Gln Glu	Ala 3940	Tyr Glu Gln Ala	Val 3945 Ala Asn Gly
Asp Ser 3950	Glu Val Val Leu	Lys 3955	Lys Leu Lys Lys	Ser 3960 Leu Asn Val
Ala Lys 3965	Ser Glu Phe Asp	Arg 3970	Asp Ala Ala Met	Gln 3975 Arg Lys Leu
Glu Lys 3980	Met Ala Asp Gln	Ala 3985	Met Thr Gln Met	Tyr 3990 Lys Gln Ala
Arg Ser 3995	Glu Asp Lys Arg	Ala 4000	Lys Val Thr Ser	Ala 4005 Met Gln Thr
Met Leu 4010	Phe Thr Met Leu	Arg 4015	Lys Leu Asp Asn	Asp 4020 Ala Leu Asn
Asn Ile 4025	Ile Asn Asn Ala	Arg 4030	Asp Gly Cys Val	Pro 4035 Leu Asn Ile
Ile Pro 4040	Leu Thr Thr Ala	Ala 4045	Lys Leu Met Val	Val 4050 Val Pro Asp
Tyr Gly 4055	Thr Tyr Lys Asn	Thr 4060	Cys Asp Gly Asn	Thr 4065 Phe Thr Tyr
Ala Ser 4070	Ala Leu Trp Glu	Ile 4075	Gln Gln Val Val	Asp 4080 Ala Asp Ser
Lys Ile 4085	Val Gln Leu Ser	Glu 4090	Ile Asn Met Asp	Asn 4095 Ser Pro Asn
Leu Ala 4100	Trp Pro Leu Ile	Val 4105	Thr Ala Leu Arg	Ala 4110 Asn Ser Ala
Val Lys 4115	Leu Gln Asn Asn	Glu 4120	Leu Ser Pro Val	Ala 4125 Leu Arg Gln
Met Ser 4130	Cys Ala Ala Gly	Thr 4135	Thr Gln Thr Ala	Cys 4140 Thr Asp Asp
Asn Ala 4145	Leu Ala Tyr Tyr	Asn 4150	Asn Ser Lys Gly	Gly 4155 Arg Phe Val
Leu Ala	Leu Leu Ser Asp His	Gln Asp Leu Lys Trp	Ala Arg Phe	

4160		4165		4170
Pro Lys Ser Asp Gly Thr	Gly Thr Ile Tyr Thr	Glu Leu Glu Pro		
4175	4180	4185		
Pro Cys Arg Phe Val Thr	Asp Thr Pro Lys Gly	Pro Lys Val Lys		
4190	4195	4200		
Tyr Leu Tyr Phe Ile Lys	Gly Leu Asn Asn Leu	Asn Arg Gly Met		
4205	4210	4215		
Val Leu Gly Ser Leu Ala	Ala Thr Val Arg Leu	Gln Ala Gly Asn		
4220	4225	4230		
Ala Thr Glu Val Pro Ala	Asn Ser Thr Val Leu	Ser Phe Cys Ala		
4235	4240	4245		
Phe Ala Val Asp Pro Ala	Lys Ala Tyr Lys Asp	Tyr Leu Ala Ser		
4250	4255	4260		
Gly Gly Gln Pro Ile Thr	Asn Cys Val Lys Met	Leu Cys Thr His		
4265	4270	4275		
Thr Gly Thr Gly Gln Ala	Ile Thr Val Thr Pro	Glu Ala Asn Met		
4280	4285	4290		
Asp Gln Glu Ser Phe Gly	Gly Ala Ser Cys Cys	Leu Tyr Cys Arg		
4295	4300	4305		
Cys His Ile Asp His Pro	Asn Pro Lys Gly Phe	Cys Asp Leu Lys		
4310	4315	4320		
Gly Lys Tyr Val Gln Ile	Pro Thr Thr Cys Ala	Asn Asp Pro Val		
4325	4330	4335		
Gly Phe Thr Leu Arg Asn	Thr Val Cys Thr Val	Cys Gly Met Trp		
4340	4345	4350		
Lys Gly Tyr Gly Cys Ser	Cys Asp Gln Leu Arg	Glu Pro Leu Met		
4355	4360	4365		
Gln Ser Ala Asp Ala Ser	Thr Phe Leu Asn Gly	Phe Ala Val		
4370	4375	4380		

<210> 75
 <211> 2695
 <212> PRT
 <213> CORONAVIRUS

<400> 75

Arg Val Cys Gly Val Ser Ala Ala Arg Leu Thr Pro Cys Gly Thr Gly
 1 5 10 15
 Thr Ser Thr Asp Val Val Tyr Arg Ala Phe Asp Ile Tyr Asn Glu Lys
 20 25 30
 Val Ala Gly Phe Ala Lys Phe Leu Lys Thr Asn Cys Cys Arg Phe Gln
 35 40 45
 Glu Lys Asp Glu Glu Gly Asn Leu Leu Asp Ser Tyr Phe Val Val Lys
 50 55 60
 Arg His Thr Met Ser Asn Tyr Gln His Glu Glu Thr Ile Tyr Asn Leu
 65 70 75 80
 Val Lys Asp Cys Pro Ala Val Ala Val His Asp Phe Phe Lys Phe Arg
 85 90 95
 Val Asp Gly Asp Met Val Pro His Ile Ser Arg Gln Arg Leu Thr Lys
 100 105 110
 Tyr Thr Met Ala Asp Leu Val Tyr Ala Leu Arg His Phe Asp Glu Gly
 115 120 125
 Asn Cys Asp Thr Leu Lys Glu Ile Leu Val Thr Tyr Asn Cys Cys Asp
 130 135 140
 Asp Asp Tyr Phe Asn Lys Lys Asp Trp Tyr Asp Phe Val Glu Asn Pro
 145 150 155 160
 Asp Ile Leu Arg Val Tyr Ala Asn Leu Gly Glu Arg Val Arg Gln Ser
 165 170 175
 Leu Leu Lys Thr Val Gln Phe Cys Asp Ala Met Arg Asp Ala Gly Ile
 180 185 190
 Val Gly Val Leu Thr Leu Asp Asn Gln Asp Leu Asn Gly Asn Trp Tyr
 195 200 205
 Asp Phe Gly Asp Phe Val Gln Val Ala Pro Gly Cys Gly Val Pro Ile
 210 215 220
 Val Asp Ser Tyr Tyr Ser Leu Leu Met Pro Ile Leu Thr Leu Thr Arg
 225 230 235 240
 Ala Leu Ala Ala Glu Ser His Met Asp Ala Asp Leu Ala Lys Pro Leu
 245 250 255
 Ile Lys Trp Asp Leu Leu Lys Tyr Asp Phe Thr Glu Glu Arg Leu Cys

260					265					270						
Leu	Phe	Asp 275	Arg	Tyr	Phe	Lys	Tyr 280	Trp	Asp	Gln	Thr	Tyr 285	His	Pro	Asn	
Cys	Ile 290	Asn	Cys	Leu	Asp	Asp 295	Arg	Cys	Ile	Leu	His 300	Cys	Ala	Asn	Phe	
Asn 305	Val	Leu	Phe	Ser	Thr 310	Val	Phe	Pro	Pro	Thr 315	Ser	Phe	Gly	Pro	Leu 320	
Val	Arg	Lys	Ile	Phe 325	Val	Asp	Gly	Val	Pro 330	Phe	Val	Val	Ser	Thr 335	Gly	
Tyr	His	Phe	Arg 340	Glu	Leu	Gly	Val	Val 345	His	Asn	Gln	Asp	Val 350	Asn	Leu	
His	Ser	Ser 355	Arg	Leu	Ser	Phe	Lys 360	Glu	Leu	Leu	Val	Tyr 365	Ala	Ala	Asp	
Pro	Ala 370	Met	His	Ala	Ala	Ser 375	Gly	Asn	Leu	Leu	Leu 380	Asp	Lys	Arg	Thr	
Thr 385	Cys	Phe	Ser	Val	Ala 390	Ala	Leu	Thr	Asn	Asn 395	Val	Ala	Phe	Gln	Thr 400	
Val	Lys	Pro	Gly	Asn 405	Phe	Asn	Lys	Asp	Phe 410	Tyr	Asp	Phe	Ala	Val 415	Ser	
Lys	Gly	Phe	Phe 420	Lys	Glu	Gly	Ser	Ser 425	Val	Glu	Leu	Lys	His 430	Phe	Phe	
Phe	Ala	Gln 435	Asp	Gly	Asn	Ala	Ala 440	Ile	Ser	Asp	Tyr	Asp 445	Tyr	Tyr	Arg	
Tyr	Asn 450	Leu	Pro	Thr	Met	Cys 455	Asp	Ile	Arg	Gln	Leu 460	Leu	Phe	Val	Val	
Glu 465	Val	Val	Asp	Lys	Tyr 470	Phe	Asp	Cys	Tyr	Asp 475	Gly	Gly	Cys	Ile	Asn 480	
Ala	Asn	Gln	Val	Ile 485	Val	Asn	Asn	Leu	Asp 490	Lys	Ser	Ala	Gly	Phe 495	Pro	
Phe	Asn	Lys	Trp 500	Gly	Lys	Ala	Arg	Leu 505	Tyr	Tyr	Asp	Ser	Met 510	Ser	Tyr	
Glu	Asp	Gln 515	Asp	Ala	Leu	Phe	Ala 520	Tyr	Thr	Lys	Arg	Asn 525	Val	Ile	Pro	
Thr	Ile	Thr	Gln	Met	Asn	Leu	Lys	Tyr	Ala	Ile	Ser	Ala	Lys	Asn	Arg	

530					535					540					
Ala 545	Arg	Thr	Val	Ala	Gly 550	Val	Ser	Ile	Cys	Ser 555	Thr	Met	Thr	Asn	Arg 560
Gln	Phe	His	Gln	Lys 565	Leu	Leu	Lys	Ser	Ile 570	Ala	Ala	Thr	Arg	Gly 575	Ala
Thr	Val	Val	Ile 580	Gly	Thr	Ser	Lys	Phe 585	Tyr	Gly	Gly	Trp	His 590	Asn	Met
Leu	Lys	Thr 595	Val	Tyr	Ser	Asp	Val 600	Glu	Thr	Pro	His	Leu 605	Met	Gly	Trp
Asp	Tyr 610	Pro	Lys	Cys	Asp	Arg 615	Ala	Met	Pro	Asn	Met 620	Leu	Arg	Ile	Met
Ala 625	Ser	Leu	Val	Leu	Ala 630	Arg	Lys	His	Asn	Thr 635	Cys	Cys	Asn	Leu	Ser 640
His	Arg	Phe	Tyr	Arg 645	Leu	Ala	Asn	Glu	Cys 650	Ala	Gln	Val	Leu	Ser 655	Glu
Met	Val	Met	Cys 660	Gly	Gly	Ser	Leu	Tyr 665	Val	Lys	Pro	Gly	Gly 670	Thr	Ser
Ser	Gly	Asp 675	Ala	Thr	Thr	Ala	Tyr 680	Ala	Asn	Ser	Val	Phe 685	Asn	Ile	Cys
Gln 690	Ala	Val	Thr	Ala	Asn	Val 695	Asn	Ala	Leu	Leu	Ser 700	Thr	Asp	Gly	Asn
Lys 705	Ile	Ala	Asp	Lys	Tyr 710	Val	Arg	Asn	Leu	Gln 715	His	Arg	Leu	Tyr	Glu 720
Cys	Leu	Tyr	Arg	Asn 725	Arg	Asp	Val	Asp	His 730	Glu	Phe	Val	Asp	Glu 735	Phe
Tyr	Ala	Tyr	Leu 740	Arg	Lys	His	Phe	Ser 745	Met	Met	Ile	Leu	Ser 750	Asp	Asp
Ala	Val	Val 755	Cys	Tyr	Asn	Ser	Asn 760	Tyr	Ala	Ala	Gln	Gly 765	Leu	Val	Ala
Ser	Ile 770	Lys	Asn	Phe	Lys	Ala 775	Val	Leu	Tyr	Tyr	Gln 780	Asn	Asn	Val	Phe
Met 785	Ser	Glu	Ala	Lys	Cys 790	Trp	Thr	Glu	Thr	Asp 795	Leu	Thr	Lys	Gly	Pro 800
His	Glu	Phe	Cys	Ser	Gln	His	Thr	Met	Leu	Val	Lys	Gln	Gly	Asp	Asp

805 810 815

Tyr Val Tyr Leu Pro Tyr Pro Asp Pro Ser Arg Ile Leu Gly Ala Gly
 820 825 830

Cys Phe Val Asp Asp Ile Val Lys Thr Asp Gly Thr Leu Met Ile Glu
 835 840 845

Arg Phe Val Ser Leu Ala Ile Asp Ala Tyr Pro Leu Thr Lys His Pro
 850 855 860

Asn Gln Glu Tyr Ala Asp Val Phe His Leu Tyr Leu Gln Tyr Ile Arg
865 870 875 880

Lys Leu His Asp Glu Leu Thr Gly His Met Leu Asp Met Tyr Ser Val
 885 890 895

Met Leu Thr Asn Asp Asn Thr Ser Arg Tyr Trp Glu Pro Glu Phe Tyr
 900 905 910

Glu Ala Met Tyr Thr Pro His Thr Val Leu Gln Ala Val Gly Ala Cys
 915 920 925

Val Leu Cys Asn Ser Gln Thr Ser Leu Arg Cys Gly Ala Cys Ile Arg
 930 935 940

Arg Pro Phe Leu Cys Cys Lys Cys Cys Tyr Asp His Val Ile Ser Thr
945 950 955 960

Ser His Lys Leu Val Leu Ser Val Asn Pro Tyr Val Cys Asn Ala Pro
 965 970 975

Gly Cys Asp Val Thr Asp Val Thr Gln Leu Tyr Leu Gly Gly Met Ser
 980 985 990

Tyr Tyr Cys Lys Ser His Lys Pro Pro Ile Ser Phe Pro Leu Cys Ala
 995 1000 1005

Asn Gly Gln Val Phe Gly Leu Tyr Lys Asn Thr Cys Val Gly Ser
1010 1015 1020

Asp Asn Val Thr Asp Phe Asn Ala Ile Ala Thr Cys Asp Trp Thr
1025 1030 1035

Asn Ala Gly Asp Tyr Ile Leu Ala Asn Thr Cys Thr Glu Arg Leu
1040 1045 1050

Lys Leu Phe Ala Ala Glu Thr Leu Lys Ala Thr Glu Glu Thr Phe
1055 1060 1065

Lys Leu Ser Tyr Gly Ile Ala Thr Val Arg Glu Val Leu Ser Asp

ES 2 396 127 T3

1070		1075		1080
Arg Glu 1085	Leu His Leu Ser	Trp 1090	Glu Val Gly Lys	Pro 1095
Arg Pro Pro				
Leu Asn 1100	Arg Asn Tyr Val	Phe 1105	Thr Gly Tyr Arg	Val 1110
Thr Lys Asn				
Ser Lys 1115	Val Gln Ile Gly	Glu 1120	Tyr Thr Phe Glu	Lys 1125
Gly Asp 1130	Ala Val Val Tyr	Arg 1135	Gly Thr Thr Thr	Tyr 1140
Lys Leu Asn				
Val Gly 1145	Asp Tyr Phe Val	Leu 1150	Thr Ser His Thr	Val 1155
Met Pro Leu				
Ser Ala 1160	Pro Thr Leu Val	Pro 1165	Gln Glu His Tyr	Val 1170
Arg Ile Thr				
Gly Leu 1175	Tyr Pro Thr Leu	Asn 1180	Ile Ser Asp Glu	Phe 1185
Ser Ser Asn				
Val Ala 1190	Asn Tyr Gln Lys	Val 1195	Gly Met Gln Lys	Tyr 1200
Ser Thr Leu				
Gln Gly 1205	Pro Pro Gly Thr	Gly 1210	Lys Ser His Phe	Ala 1215
Ile Gly Leu				
Ala Leu 1220	Tyr Tyr Pro Ser	Ala 1225	Arg Ile Val Tyr	Thr 1230
Ala Cys Ser				
His Ala 1235	Ala Val Asp Ala	Leu 1240	Cys Glu Lys Ala	Leu 1245
Lys Tyr Leu				
Pro Ile 1250	Asp Lys Cys Ser	Arg 1255	Ile Ile Pro Ala	Arg 1260
Ala Arg Val				
Glu Cys 1265	Phe Asp Lys Phe	Lys 1270	Val Asn Ser Thr	Leu 1275
Glu Gln Tyr				
Val Phe 1280	Cys Thr Val Asn	Ala 1285	Leu Pro Glu Thr	Thr 1290
Ala Asp Ile				
Val Val 1295	Phe Asp Glu Ile	Ser 1300	Met Ala Thr Asn	Tyr 1305
Asp Leu Ser				
Val Val 1310	Asn Ala Arg Leu	Arg 1315	Ala Lys His Tyr	Val 1320
Tyr Ile Gly				
Asp Pro	Ala Gln Leu Pro Ala	Pro Arg Thr Leu Leu	Thr Lys Gly	

1325						1330						1335			
Thr	Leu 1340	Glu	Pro	Glu	Tyr	Phe 1345	Asn	Ser	Val	Cys	Arg 1350	Leu	Met	Lys	
Thr	Ile 1355	Gly	Pro	Asp	Met	Phe 1360	Leu	Gly	Thr	Cys	Arg 1365	Arg	Cys	Pro	
Ala	Glu 1370	Ile	Val	Asp	Thr	Val 1375	Ser	Ala	Leu	Val	Tyr 1380	Asp	Asn	Lys	
Leu	Lys 1385	Ala	His	Lys	Asp	Lys 1390	Ser	Ala	Gln	Cys	Phe 1395	Lys	Met	Phe	
Tyr	Lys 1400	Gly	Val	Ile	Thr	His 1405	Asp	Val	Ser	Ser	Ala 1410	Ile	Asn	Arg	
Pro	Gln 1415	Ile	Gly	Val	Val	Arg 1420	Glu	Phe	Leu	Thr	Arg 1425	Asn	Pro	Ala	
Trp	Arg 1430	Lys	Ala	Val	Phe	Ile 1435	Ser	Pro	Tyr	Asn	Ser 1440	Gln	Asn	Ala	
Val	Ala 1445	Ser	Lys	Ile	Leu	Gly 1450	Leu	Pro	Thr	Gln	Thr 1455	Val	Asp	Ser	
Ser	Gln 1460	Gly	Ser	Glu	Tyr	Asp 1465	Tyr	Val	Ile	Phe	Thr 1470	Gln	Thr	Thr	
Glu	Thr 1475	Ala	His	Ser	Cys	Asn 1480	Val	Asn	Arg	Phe	Asn 1485	Val	Ala	Ile	
Thr	Arg 1490	Ala	Lys	Ile	Gly	Ile 1495	Leu	Cys	Ile	Met	Ser 1500	Asp	Arg	Asp	
Leu	Tyr 1505	Asp	Lys	Leu	Gln	Phe 1510	Thr	Ser	Leu	Glu	Ile 1515	Pro	Arg	Arg	
Asn	Val 1520	Ala	Thr	Leu	Gln	Ala 1525	Glu	Asn	Val	Thr	Gly 1530	Leu	Phe	Lys	
Asp	Cys 1535	Ser	Lys	Ile	Ile	Thr 1540	Gly	Leu	His	Pro	Thr 1545	Gln	Ala	Pro	
Thr	His 1550	Leu	Ser	Val	Asp	Ile 1555	Lys	Phe	Lys	Thr	Glu 1560	Gly	Leu	Cys	
Val	Asp 1565	Ile	Pro	Gly	Ile	Pro 1570	Lys	Asp	Met	Thr	Tyr 1575	Arg	Arg	Leu	
Ile	Ser	Met	Met	Gly	Phe	Lys	Met	Asn	Tyr	Gln	Val	Asn	Gly	Tyr	

1580		1585		1590
Pro Asn 1595	Met Phe Ile Thr	Arg 1600	Glu Glu Ala Ile	Arg 1605 His Val Arg
Ala Trp 1610	Ile Gly Phe Asp	Val 1615	Glu Gly Cys His	Ala 1620 Thr Arg Asp
Ala Val 1625	Gly Thr Asn Leu	Pro 1630	Leu Gln Leu Gly	Phe 1635 Ser Thr Gly
Val Asn 1640	Leu Val Ala Val	Pro 1645	Thr Gly Tyr Val	Asp 1650 Thr Glu Asn
Asn Thr 1655	Glu Phe Thr Arg	Val 1660	Asn Ala Lys Pro	Pro 1665 Pro Gly Asp
Gln Phe 1670	Lys His Leu Ile	Pro 1675	Leu Met Tyr Lys	Gly 1680 Leu Pro Trp
Asn Val 1685	Val Arg Ile Lys	Ile 1690	Val Gln Met Leu	Ser 1695 Asp Thr Leu
Lys Gly 1700	Leu Ser Asp Arg	Val 1705	Val Phe Val Leu	Trp 1710 Ala His Gly
Phe Glu 1715	Leu Thr Ser Met	Lys 1720	Tyr Phe Val Lys	Ile 1725 Gly Pro Glu
Arg Thr 1730	Cys Cys Leu Cys	Asp 1735	Lys Arg Ala Thr	Cys 1740 Phe Ser Thr
Ser Ser 1745	Asp Thr Tyr Ala	Cys 1750	Trp Asn His Ser	Val 1755 Gly Phe Asp
Tyr Val 1760	Tyr Asn Pro Phe	Met 1765	Ile Asp Val Gln	Gln 1770 Trp Gly Phe
Thr Gly 1775	Asn Leu Gln Ser	Asn 1780	His Asp Gln His	Cys 1785 Gln Val His
Gly Asn 1790	Ala His Val Ala	Ser 1795	Cys Asp Ala Ile	Met 1800 Thr Arg Cys
Leu Ala 1805	Val His Glu Cys	Phe 1810	Val Lys Arg Val	Asp 1815 Trp Ser Val
Glu Tyr 1820	Pro Ile Ile Gly	Asp 1825	Glu Leu Arg Val	Asn 1830 Ser Ala Cys
Arg Lys	Val Gln His Met Val	Val Lys Ser Ala	Leu	Leu Ala Asp

1835					1840					1845				
Lys Phe	Pro Val	Leu His	Asp	Ile Gly	Asn Pro	Lys	Ala Ile	Lys						
1850			1855			1860								
Cys Val	Pro Gln	Ala Glu	Val	Glu Trp	Lys Phe	Tyr	Asp Ala	Gln						
1865			1870			1875								
Pro Cys	Ser Asp	Lys Ala	Tyr	Lys Ile	Glu Glu	Leu	Phe Tyr	Ser						
1880			1885			1890								
Tyr Ala	Thr His	His Asp	Lys	Phe Thr	Asp Gly	Val	Cys Leu	Phe						
1895			1900			1905								
Trp Asn	Cys Asn	Val Asp	Arg	Tyr Pro	Ala Asn	Ala	Ile Val	Cys						
1910			1915			1920								
Arg Phe	Asp Thr	Arg Val	Leu	Ser Asn	Leu Asn	Leu	Pro Gly	Cys						
1925			1930			1935								
Asp Gly	Gly Ser	Leu Tyr	Val	Asn Lys	His Ala	Phe	His Thr	Pro						
1940			1945			1950								
Ala Phe	Asp Lys	Ser Ala	Phe	Thr Asn	Leu Lys	Gln	Leu Pro	Phe						
1955			1960			1965								
Phe Tyr	Tyr Ser	Asp Ser	Pro	Cys Glu	Ser His	Gly	Lys Gln	Val						
1970			1975			1980								
Val Ser	Asp Ile	Asp Tyr	Val	Pro Leu	Lys Ser	Ala	Thr Cys	Ile						
1985			1990			1995								
Thr Arg	Cys Asn	Leu Gly	Gly	Ala Val	Cys Arg	His	His Ala	Asn						
2000			2005			2010								
Glu Tyr	Arg Gln	Tyr Leu	Asp	Ala Tyr	Asn Met	Met	Ile Ser	Ala						
2015			2020			2025								
Gly Phe	Ser Leu	Trp Ile	Tyr	Lys Gln	Phe Asp	Thr	Tyr Asn	Leu						
2030			2035			2040								
Trp Asn	Thr Phe	Thr Arg	Leu	Gln Ser	Leu Glu	Asn	Val Ala	Tyr						
2045			2050			2055								
Asn Val	Val Asn	Lys Gly	His	Phe Asp	Gly His	Ala	Gly Glu	Ala						
2060			2065			2070								
Pro Val	Ser Ile	Ile Asn	Asn	Ala Val	Tyr Thr	Lys	Val Asp	Gly						
2075			2080			2085								
Ile Asp	Val Glu	Ile Phe	Glu	Asn Lys	Thr Thr	Leu	Pro Val	Asn						

ES 2 396 127 T3

2090	2095	2100
Val Ala Phe Glu Leu Trp	Ala Lys Arg Asn Ile	Lys Pro Val Pro
2105	2110	2115
Glu Ile Lys Ile Leu Asn	Asn Leu Gly Val Asp	Ile Ala Ala Asn
2120	2125	2130
Thr Val Ile Trp Asp Tyr	Lys Arg Glu Ala Pro	Ala His Val Ser
2135	2140	2145
Thr Ile Gly Val Cys Thr	Met Thr Asp Ile Ala	Lys Lys Pro Thr
2150	2155	2160
Glu Ser Ala Cys Ser Ser	Leu Thr Val Leu Phe	Asp Gly Arg Val
2165	2170	2175
Glu Gly Gln Val Asp Leu	Phe Arg Asn Ala Arg	Asn Gly Val Leu
2180	2185	2190
Ile Thr Glu Gly Ser Val	Lys Gly Leu Thr Pro	Ser Lys Gly Pro
2195	2200	2205
Ala Gln Ala Ser Val Asn	Gly Val Thr Leu Ile	Gly Glu Ser Val
2210	2215	2220
Lys Thr Gln Phe Asn Tyr	Phe Lys Lys Val Asp	Gly Ile Ile Gln
2225	2230	2235
Gln Leu Pro Glu Thr Tyr	Phe Thr Gln Ser Arg	Asp Leu Glu Asp
2240	2245	2250
Phe Lys Pro Arg Ser Gln	Met Glu Thr Asp Phe	Leu Glu Leu Ala
2255	2260	2265
Met Asp Glu Phe Ile Gln	Arg Tyr Lys Leu Glu	Gly Tyr Ala Phe
2270	2275	2280
Glu His Ile Val Tyr Gly	Asp Phe Ser His Gly	Gln Leu Gly Gly
2285	2290	2295
Leu His Leu Met Ile Gly	Leu Ala Lys Arg Ser	Gln Asp Ser Pro
2300	2305	2310
Leu Lys Leu Glu Asp Phe	Ile Pro Met Asp Ser	Thr Val Lys Asn
2315	2320	2325
Tyr Phe Ile Thr Asp Ala	Gln Thr Gly Ser Ser	Lys Cys Val Cys
2330	2335	2340
Ser Val Ile Asp Leu Leu Leu	Asp Asp Phe Val Glu	Ile Ile Lys

2345		2350		2355
Ser Gln 2360	Asp Leu Ser Val	Ile 2365	Ser Lys Val Val	Lys 2370 Val Thr Ile
Asp Tyr 2375	Ala Glu Ile Ser	Phe 2380	Met Leu Trp Cys	Lys 2385 Asp Gly His
Val Glu 2390	Thr Phe Tyr Pro	Lys 2395	Leu Gln Ala Ser	Gln 2400 Ala Trp Gln
Pro Gly 2405	Val Ala Met Pro	Asn 2410	Leu Tyr Lys Met	Gln 2415 Arg Met Leu
Leu Glu 2420	Lys Cys Asp Leu	Gln 2425	Asn Tyr Gly Glu	Asn 2430 Ala Val Ile
Pro Lys 2435	Gly Ile Met Met	Asn 2440	Val Ala Lys Tyr	Thr 2445 Gln Leu Cys
Gln Tyr 2450	Leu Asn Thr Leu	Thr 2455	Leu Ala Val Pro	Tyr 2460 Asn Met Arg
Val Ile 2465	His Phe Gly Ala	Gly 2470	Ser Asp Lys Gly	Val 2475 Ala Pro Gly
Thr Ala 2480	Val Leu Arg Gln	Trp 2485	Leu Pro Thr Gly	Thr 2490 Leu Leu Val
Asp Ser 2495	Asp Leu Asn Asp	Phe 2500	Val Ser Asp Ala	Asp 2505 Ser Thr Leu
Ile Gly 2510	Asp Cys Ala Thr	Val 2515	His Thr Ala Asn	Lys 2520 Trp Asp Leu
Ile Ile 2525	Ser Asp Met Tyr	Asp 2530	Pro Arg Thr Lys	His 2535 Val Thr Lys
Glu Asn 2540	Asp Ser Lys Glu	Gly 2545	Phe Phe Thr Tyr	Leu 2550 Cys Gly Phe
Ile Lys 2555	Gln Lys Leu Ala	Leu 2560	Gly Gly Ser Ile	Ala 2565 Val Lys Ile
Thr Glu 2570	His Ser Trp Asn	Ala 2575	Asp Leu Tyr Lys	Leu 2580 Met Gly His
Phe Ser 2585	Trp Trp Thr Ala	Phe 2590	Val Thr Asn Val	Asn 2595 Ala Ser Ser
Ser Glu	Ala Phe Leu Ile Gly	Ala Asn Tyr Leu Gly	Lys Pro Lys	

2600		2605		2610
Glu Gln Ile Asp Gly Tyr Thr	Met His Ala Asn Tyr	Ile Phe Trp		
2615	2620	2625		
Arg Asn Thr Asn Pro Ile Gln	Leu Ser Ser Tyr Ser	Leu Phe Asp		
2630	2635	2640		
Met Ser Lys Phe Pro Leu Lys	Leu Arg Gly Thr Ala	Val Met Ser		
2645	2650	2655		
Leu Lys Glu Asn Gln Ile Asn	Asp Met Ile Tyr Ser	Leu Leu Glu		
2660	2665	2670		
Lys Gly Arg Leu Ile Ile Arg	Glu Asn Asn Arg Val	Val Val Ser		
2675	2680	2685		
Ser Asp Ile Leu Val Asn Asn				
2690	2695			

<210> 76
 <211> 20
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> cebador S/L3/+4932

<400> 76
 ccacacacag cttgttgata 20

<210> 77
 <211> 20
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> cebador S/L4/+6401

<400> 77
 ccgaagttgt aggcaatgtc 20

<210> 78
 <211> 20
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> cebador S/L4/+6964

<400> 78 20
 ttggtgctc cttctattg 20

<210> 79
 <211> 20
 <212> ADN
 <213> Secuencia artificial

<220>

<223> cebador S/L4/-/6817

<400> 79 20

ccggcatcca aacataattt

20

<210> 80

<211> 20

<212> ADN

<213> Secuencia artificial

<220>

<223> cebador S/L5/-/7633

<400> 80

tggtcagtag ggttgattgg

20

<210> 81

<211> 20

<212> ADN

<213> Secuencia artificial

<220>

<223> S/L5/-/8127

<400> 81

catcctttgt gtcaacatcg

20

<210> 82

<211> 20

<212> ADN

<213> Secuencia artificial

<220>

<223> cebador S/L5/-/8633

<400> 82

gtcacgagtg acaccatcct

20

<210> 83

<211> 20

<212> ADN

<213> Secuencia artificial

<220>

<223> cebador S/L5/+7839

<400> 83

atgacgacgag tctgcttcta

20

<210> 84

<211> 20

<212> ADN

<213> Secuencia artificial

<220>

<223> cebador S/L5/+7885

<400> 84

ttcatagtag ctggcttacc

20

<210> 85

<211> 20

<212> ADN

<213> Secuencia artificial

<220>

<223> cebador S/L5/+8255

<400> 85
atcttgccgc atgtattgac 20

<210> 86
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador S/L6/-9422

<400> 86
tgcattagca gcaacaacat 20

<210> 87
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador 5/L6/-9966

<400> 87
tctgcagaac agcagaagtg 20

<210> 88
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador S/L6/-10542

<400> 88
cctgtgcagt ttgtctgtca 20

<210> 89
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador S/L6/+10677

<400> 89
ccttggtgca atgaagtaca 20

<210> 90
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador S/L6/+10106

<400> 90
atgtcatttg cacagcagaa 20

<210> 91
<211> 20
<212> ADN
<213> Secuencia artificial

<220>

<223> cebador S/L6/+ /9571

<400> 91

cttcaatggt ttgccatgtt

20

<210> 92

<211> 20

<212> ADN

<213> Secuencia artificial

<220>

<223> cebador S/L7/- /11271

<400> 92

tgcgagctgt catgagaata

20

<210> 93

<211> 20

<212> ADN

<213> Secuencia artificial

<220>

<223> cebador S/L7/- /11801

<400> 93

aaccgagagc agtaccacag

20

<210> 94

<211> 20

<212> ADN

<213> Secuencia artificial

<220>

<223> cebador S/L7/- /12383

<400> 94

tttggtgct gtagtcaatg

20

<210> 95

<211> 20

<212> ADN

<213> Secuencia artificial

<220>

<223> cebador S/L7/+ /12640

<400> 95

ctacgacaga tgcctgtgc

20

<210> 96

<211> 20

<212> ADN

<213> Secuencia artificial

<220>

<223> cebador S/L7/+ /12088

<400> 96

gagcaggctg tagctaattg

20

<210> 97

<211> 20

<212> ADN

<213> Secuencia artificial

<220>

<223> cebador S/L7/+11551

<400> 97

ttaggctatt gttgctgctg

20

<210> 98

<211> 20

<212> ADN

<213> Secuencia artificial

<220>

<223> cebador S/L8/-13160

<400> 98

cagacaacat gaagcaccac

20

<210> 99

<211> 20

<212> ADN

<213> Secuencia artificial

<220>

<223> cebador S/L8/-13704

<400> 99

cgctgacgtg atatgttg

20

<210> 100

<211> 20

<212> ADN

<213> Secuencia artificial

<220>

<223> cebador S/L8/-14284

<400> 100

tgcaaatga aggatacacc

20

<210> 101

<211> 20

<212> ADN

<213> Secuencia artificial

<220>

<223> cebador S/L8/+14453

<400> 101

acatagctcg cgtctcagtt

20

<210> 102

<211> 20

<212> ADN

<213> Secuencia artificial

<220>

<223> cebador S/L8/+13968

<400> 102

ggcattgtag gcgtactgac

20

<210> 103

<211> 19

<212> ADN

<213> Secuencia artificial

<220>

<223> cebador S/L8/+13401

<400> 103
gtttgcggtg taagtgcag 19

<210> 104
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador S/L9/-15098

<400> 104
tagtggcggc tattgacttc 20

<210> 105
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador S/L9/-15677

<400> 105
ctaaacctg agccgcatag 20

<210> 106
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador S/L9/-16247

<400> 106
catggtcata gcagcacttg 20

<210> 107
<211> 21
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador S/L9/+16323

<400> 107
ccaggttgat atgtcactga t 21

<210> 108
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador S/L9/+15858

<400> 108
ccttaccag atccatcaag 20

<210> 109
<211> 20
<212> ADN
<213> Secuencia artificial

<220>

<223> cebador S/L9/+15288

<400> 109
cgcaaacata acacttgctg 20

<210> 110
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador S/L10/-16914

<400> 110
agtgttggt acaagccagt 20

<210> 111
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador S/L10/-17466

<400> 111
gtccaagga acatgtctg 20

<210> 112
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador S/L10/-18022

<400> 112
agggtcctgt gtaggatgaa 20

<210> 113
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador S/L10/+18245

<400> 113
gggctgtcat gcaactagag 20

<210> 114
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador S/L10/+17663

<400> 114
tcttacacgc aatcctgctt 20

<210> 115
<211> 20
<212> ADN
<213> Secuencia artificial

<220>

<223> cebador S/L10/+17061

<400> 115
taccatctg ctgcatagt 20

<210> 116
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador S/L11/-18877

<400> 116
gcaagcagaa ttaaccctca 20

<210> 117
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador S/L11/-19396

<400> 117
agcaccacct aaattgcatc 20

<210> 118
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador S/L11/-20002

<400> 118
tggcccttt gaaggtgta 20

<210> 119
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador S/L11/+20245

<400> 119
tcgaacacat cgtttatgga 20

<210> 120
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador S/L11/+19611

<400> 120
gaagcacctg ttccatcat 20

<210> 121
<211> 20
<212> ADN
<213> Secuencia artificial

<220>

<223> cebador S/L11/+19021

<400> 121
acgatgctca gccatgtagt 20

<210> 122
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador SRAS/L1/F3/+800

<400> 122
gaggtgcagt cactcgctat 20

<210> 123
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador SRAS/L1/F4/+1391

<400> 123
cagagattgg acctgagcat 20

<210> 124
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador SRAS/L1/F5/+1925

<400> 124
cagcaaacca ctcaattcct 20

<210> 125
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador SRAS/L1/R3/-1674

<400> 125
aaatgatggc aacctcttca 20

<210> 126
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador SRAS/L1/R4/-1107

<400> 126
cacgtggttg aatgactttg 20

<210> 127
<211> 20
<212> ADN
<213> Secuencia artificial

<220>

<223> cebador SRAS/L1/R5/-/520

<400> 127
atttctgcaa ccagctcaac 20

<210> 128
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador SRAS/L2/F3/+ /2664

<400> 128
cgcatgtgtc cctggtttac 20

<210> 129
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador SRAS/L2/F4/+ /3232

<400> 129
gagattgagc cagaaccaga 20

<210> 130
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador SRAS/L2/F5/+ /3746

<400> 130
atgagcaggt tgtcatggat 20

<210> 131
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador SRAS/L2/R3/- /3579

<400> 131
ctgccttaag aagctggatg 20

<210> 132
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador SRAS/L2/R4/- /2991

<400> 132
tttcttcacc agcatcatca 20

<210> 133
<211> 20
<212> ADN
<213> Secuencia artificial

<220>

<223> cebador SRAS/L2/R5/-/2529

<400> 133
caccgttctt gagaacaacc 20

<210> 134
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador SRAS/L3/F3/+4708

<400> 134
tctttggctg gctcttacag 20

<210> 135
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador SRAS/L3/F4/+5305

<400> 135
gctggtgatg ctgctaactt 20

<210> 136
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador SRAS/L3/F5/+5822

<400> 136
ccatcaagcc tgtgtcgtat 20

<210> 137
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador SRAS/L3/R3/-/5610

<400> 137
caggtggtgc agacatcata 20

<210> 138
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador SRAS/L3/R4/-/4988

<400> 138
aacatcagca ccatccaagt 20

<210> 139
<211> 20
<212> ADN
<213> Secuencia artificial

<220>

<223> cebador SRAS/L3/R5/-/4437

<400> 139

atcggacacc atagtcaacg

20

<210> 140

<211> 7788

<212> ADN

<213> Secuencia artificial

<220>

<223> gen S sintético

<400> 140

tcaatattgg ccattagcca tattattcat tggttatata gcataaatca atattggcta	60
ttggccattg catagcttgt atctatatca taatatgtac atttatattg gtcattgtcc	120
aatatgaccg ccatgttggc attgattatt gactagttat taatagtaat caattacggg	180
gtcattagtt catagcccat atatggagtt ccgcgttaca taacttacgg taaatggccc	240
gcctggctga ccgccaacg acccccgcgc attgacgtca ataatgacgt atgttcccat	300
agtaacgcca atagggactt tccattgacg tcaatgggtg gagtatttac ggtaaactgc	360
ccacttggca gtacatcaag tgtatcatat gccaagtccg cccctattg acgtcaatga	420
cggtaaattg cccgcctggc attatgcca gtacatgacc ttacgggact ttcctacttg	480
gcagtacatc tacgtattag tcatcgctat taccatgggt atgcgggtttt ggcagtacac	540
caatgggagc ggatagcggc ttgactcacg gggatttcca agtctccacc ccattgacgt	600
caatgggagc ttgttttggc accaaaatca acgggacttt ccaaaatgtc gtaataaccc	660
cgccccgttg acgcaaattg gcggtaggcg tgtacgggtg gaggtctata taagcagagc	720
tcgttttagt aaccgtcaga tcactagaag ctttattgct gtagtttatc acagttaaat	780
tgctaacgca gtcagtgtt ctgacacaac agtctcgaac ttaagctgca gaagttggc	840
gtgaggcact gggcaggtaa gtatcaaggt tacaagacag gtttaaggag accaatagaa	900

actgggcttg	tcgagacaga	gaagactctt	gcgtttctga	taggcaccta	ttggtcttac	960
tgacatccac	tttgcctttc	tctccacagg	tgtccactcc	cagttcaatt	acagctctta	1020
aggctagagt	acttaatacg	actcactata	ggctagcggg	tccaccatgt	tcatcttcct	1080
gctgttcctg	accctgacca	gcggcagcga	cctggaccgg	tgcaccacct	tcgacgacgt	1140
gcaggcccc	aactacaccc	agcacaccag	cagcatgcgg	ggcgtgtact	accccgacga	1200
gatctttcgg	agcgacaccc	tgtacctgac	ccaggacctg	ttcctgccct	tctacagcaa	1260
cgtgaccggc	ttccacacca	tcaaccacac	cttcggcaac	cccgtgatcc	ccttcaagga	1320
cggcatctac	ttcgccgcca	ccgagaagag	caacgtgggtg	cggggctggg	tgttcggcag	1380
caccatgaac	aacaagagcc	agagcgtgat	catcatcaac	aacagcacca	acgtggtgat	1440
ccgggcctgc	aacttcgagc	tgtgcgacaa	ccccttcttc	gccgtgtcca	aacctatggg	1500
caccagacc	cacaccatga	tcttcgacaa	cgccttcaac	tgcaccttcg	agtacatcag	1560
cgacgccttc	agcctggacg	tgagcgagaa	gagcggcaac	ttcaagcacc	tgcgggagtt	1620
cgtgttcaag	aacaaggacg	gcttcctgta	cgtgtacaag	ggctaccagc	ccatcgacgt	1680
ggtgagagac	ctgcccagcg	gcttcaacac	cctgaagccc	atcttcaagc	tgcccctggg	1740
catcaacatc	accaacttcc	gggccatcct	gaccgccttt	agccctgccc	aggacatctg	1800
gggcaccagc	gccgccgcct	acttcgtggg	ctacctgaag	cctaccacct	tcatgctgaa	1860
gtacgacgag	aacggcacca	tcaccgacgc	cgtggactgc	agccagaacc	ccctggccga	1920
gctgaagtgc	agcgtgaaga	gcttcgagat	cgacaagggc	atctaccaga	ccagcaactt	1980
cagagtgggtg	cctagcggcg	atgtggtgcg	gttccccaat	atcaccaacc	tgtgcccctt	2040
cggcgaagtg	ttcaacgcca	ccaagttccc	cagcgtgtac	gcctgggagc	ggaagaagat	2100
cagcaactgc	gtggccgact	acagcgtgct	gtacaactcc	accttcttca	gcaccttcaa	2160
gtgctacggc	gtgagcgcca	ccaagctgaa	cgacctgtgc	ttcagcaacg	tgtacgccga	2220
cagcttcgtg	gtgaagggcg	acgacgtgag	acagatcgcc	cctggccaga	ccggcgtgat	2280
cgccgactac	aactacaagc	tgcccgacga	cttcatgggc	tgcgtgctgg	cctggaacac	2340
ccggaacatc	gacgccacaa	gcaccggcaa	ctacaattac	aagtaccgct	acctgcggca	2400
cggcaagctg	cggcccttcg	agcgggacat	ctccaacgtg	cccttcagcc	ccgacggcaa	2460
gccctgcacc	ccccctgccc	tgaactgcta	ctggccccctg	aacgactacg	gcttctacac	2520
caccaccggc	atcggtatc	agccctacag	agtgggtggg	ctgagcttcg	agctgctgaa	2580
cgcccctgcc	accgtgtgcg	gccccaaagct	gagcaccgac	ctgatcaaga	accagtgcgt	2640
gaacttcaac	ttcaacggcc	tgaccggcac	cggcgtgctg	acccccagca	gcaagcgctt	2700
ccagcccttc	cagcagttcg	gccgggatgt	gagcgacttc	accgacagcg	tgcgggaccc	2760
caagaccagc	gagatcctgg	acatcagccc	ctgcagcttc	ggcggcgtgt	ccgtgatcac	2820
ccccggcacc	aacgccagca	gcgaagtggc	cgtgctgtac	caggacgtga	actgcaccga	2880
cgtgagcacc	gccatccacg	ccgaccagct	gacccccgcc	tggcggatct	acagcaccgg	2940

gaacaacgtg	ttccagaccc	aggccggctg	cctgatcggc	gccgagcacg	tggacaccag	3000
ctacgagtgc	gacatcccca	ttggcgccgg	aatctgcgcc	agctaccaca	ccgtgagcct	3060
gctgcggagc	accagccaga	agtccatcgt	ggcctacacc	atgagcctgg	gcgccgacag	3120
cagcatcgcc	tacagcaaca	acaccatcgc	catccccacc	aacttcagca	tctccatcac	3180
caccgaagtg	atgcccgtga	gcatggccaa	gacaagcgtg	gattgcaaca	tgtacatctg	3240
cggcgacagc	accgagtgcg	ccaacctgct	gctgcagtac	ggcagcttct	gcacccagct	3300
gaaccggggc	ctgagcggca	tcgccgcccga	gcaggaccgg	aacaccagag	aagtgttcgc	3360
ccaagtgaag	cagatgtata	agacccccac	cctgaagtac	ttcgggggct	tcaacttctc	3420
tcagatcctg	cccgaacctc	tgaagcccac	caagcgctcc	ttcatcgagg	acctgctgtt	3480
caacaaagtg	accctggccg	acgccggctt	tatgaagcag	tacggcgagt	gcctgggcga	3540
catcaacgcc	cgggacctga	tctgcgcccga	gaagttaa	gggctgaccg	tgtgcccc	3600
cctgctgacc	gacgacatga	tcgccgccta	tacagccgcc	ctggtgagcg	gcaccgccac	3660
cgccggctgg	accttcggag	cgggagccgc	cctgcagatc	cccttcgcca	tgcagatggc	3720
ctaccggttc	aacggcatcg	gcgtgaccca	gaacgtgctg	tacgagaacc	agaagcagat	3780
cgccaaccag	ttcaacaagg	ccatcagcca	gatccaggag	agcctgacca	caaccagcac	3840
cggcctgggc	aagctgcagg	acgtggtgaa	ccagaacgcc	caggccctga	acaccctggt	3900
gaagcagctg	agcagcaact	tcggcgccat	cagctctgtg	ctgaacgaca	tcctgagcag	3960
gctggacaaa	gtggaggccg	aagtgcagat	cgaccggctg	atcaccggac	gcctgcagtc	4020
cctgcagacc	tacgtgaccc	agcagctgat	cagagccgcc	gagatccggg	ccagcgccaa	4080
tctggccgcc	accaagatga	gcgagtgcgt	gctgggccag	agcaagagag	tggacttctg	4140
cggcaagggc	tatcacctga	tgagcttccc	ccaggccgcc	ccccacggcg	tggtgttcct	4200
gcacgtgacc	tacgtgccta	gccaggagcg	gaacttcacc	accgccccag	ccatctgccca	4260
cgagggcaag	gcctacttcc	cccgggaggg	cgtgttcgtg	tttaacggca	ccagctgggt	4320
catcacccag	cgcaacttct	tcagcccca	gatcatcacc	acagacaaca	ccttcgtgtc	4380
cggcaactgt	gatgtggtga	tcggcatcat	caataacacc	gtgtacgacc	ccctgcagcc	4440
cgagctggac	agcttcaagg	aggagctgga	caaatacttc	aagaaccaca	cctccccga	4500
cgtggacctg	ggcgatatca	gcggcatcaa	cgctccgtg	gtgaacatcc	agaaggagat	4560
cgacagactg	aacgaagtgg	ccaagaacct	gaacgagagc	ctgatcgacc	tgcaggagct	4620
gggcaagtac	gagcagtaca	tcaagtggcc	ctggtacgtg	tggctgggct	tcatcgccgg	4680
cctgatcgcc	atcgtgatgg	tgaccatcct	gctgtgctgc	atgaccagct	gctgtagctg	4740
cctgaaaggc	gcctgcagct	gtggcagctg	ctgcaagttc	gacgaggacg	acagcgagcc	4800
cgtgctgaag	ggcgtgaagc	tgcactacac	ctgataactc	gagaattcac	gcgtgggtacc	4860
tctagagtgc	acccgggccc	ccgcttcgag	cagacatgat	aagatacatt	gatgagtttg	4920
gacaaaccac	aactagaatg	cagtgaaaaa	aatgctttat	ttgtgaaatt	tgtgatgcta	4980

ttgcttttatt	tgtaaccatt	ataagctgca	ataaacaagt	taacaacaac	aattgcattc	5040
attttatgtt	tcaggttcag	ggggagatgt	gggaggtttt	ttaaagcaag	taaaacctct	5100
acaaatgttg	taaaatcgat	aaggatccgg	gctggcgtaa	tagcgaagag	gcccgcaccg	5160
atcgcccttc	ccaacagttg	cgcagcctga	atggcgaatg	gacgcgccct	gtagcggcgc	5220
attaagcgcg	gcgggtgttg	tggttacgcg	cagcgtgacc	gctacacttg	ccagcgcctt	5280
agcgcccgt	cctttcgctt	tcttcccttc	ctttctcgcc	acgttcgccg	gctttccccg	5340
tcaagctcta	aatcgggggc	tccctttagg	gttccgattt	agagctttac	ggcacctcga	5400
ccgcaaaaaa	cttgatttgg	gtgatggttc	acgtagtggg	ccatcgccct	gatagacggt	5460
ttttcgccct	ttgacgttgg	agtccacgtt	ctttaatagt	ggactcttgt	tccaaactgg	5520
aacaacactc	aaccctatct	cggctctattc	ttttgattta	taagggattt	tgccgatttc	5580
ggcctatttg	ttaaaaaatg	agctgattta	acaaatattt	aacgcgaatt	ttaacaaaat	5640
attaacgttt	acaatttcgc	ctgatgcggt	attttctcct	tacgcatctg	tgcggtattt	5700
cacaccgcat	atggtgcact	ctcagtacaa	tctgctctga	tgccgcatag	ttaagccagc	5760
cccgacaccc	gccaacaccc	gctgacgcgc	cctgacgggc	ttgtctgctc	ccggcatccg	5820
cttacagaca	agctgtgacc	gtctccggga	gctgcatgtg	tcagaggttt	tcaccgtcat	5880
caccgaaacg	cgcgagacga	aagggcctcg	tgatacgcct	atttttatag	gttaatgtca	5940
tgataataat	ggtttcttag	acgtcagggtg	gcacttttctg	gggaaatgtg	cgcggaaccc	6000
ctatttgttt	atttttctaa	atacattcaa	atatgtatcc	gctcatgaga	caataacctt	6060
gataaatgct	tcaataatat	tgaaaaagga	agagtatgag	tattcaacat	ttccgtgtcg	6120
cccttattcc	cttttttgcg	gcattttgcc	ttcctgtttt	tgctcaccca	gaaacgctgg	6180
tgaaagtaaa	agatgctgaa	gatcagttgg	gtgcacgagt	gggttacatc	gaactggatc	6240
tcaacagcgg	taagatcctt	gagagttttc	gccccgaaga	acgttttcca	atgatgagca	6300
cttttaagt	tctgctatgt	ggcgcggtat	tatcccgtat	tgacgccggg	caagagcaac	6360
tcggtcgccg	catacactat	tctcagaatg	acttggttga	gtactcacca	gtcacagaaa	6420
agcatcttac	ggatggcatg	acagtaagag	aattatgcag	tgctgccata	accatgagtg	6480
ataacactgc	ggccaactta	cttctgacaa	cgatcgaggg	accgaaggag	ctaaccgctt	6540
ttttgcacaa	catgggggat	catgtaactc	gccttgatcg	ttgggaaccg	gagctgaatg	6600
aagccatacc	aaacgacgag	cgtgacacca	cgatgcctgt	agcaatggca	acaacgttgc	6660
gcaaactatt	aactggcgaa	ctacttactc	tagcttcccc	gcaacaatta	atagactgga	6720
tggaggcgga	taaagttgca	ggaccacttc	tgcgctcggc	ccttccggct	ggctggttta	6780
ttgctgataa	atctggagcc	ggtgagcgtg	ggtctcgcg	tatcattgca	gactggggc	6840
cagatggtaa	gccctcccgt	atcgtagtta	tctacacgac	ggggagtcag	gcaactatgg	6900
atgaacgaaa	tagacagatc	gctgagatag	gtgcctcact	gattaagcat	tggttaactgt	6960
cagaccaagt	ttactcatat	atactttaga	ttgatttaaa	acttcatttt	taatttaaaa	7020

ggatctaggt gaagatCctt tttgataatc tcatgaccaa aatcccttaa cgtgagtttt	7080
cgttccactg agcgtcagac cccgtagaaa agatcaaagg atcttcttga gatccttttt	7140
ttctgcgcgt aatctgCtgc ttgcaaacaa aaaaaccacc gctaccagcg gtgggtttgtt	7200
tgccggatca agagctacca actctttttt cgaaggtaac tggcttcagc agagcgcaga	7260
taccaaatac tgtccttcta gtgtagccgt agttaggcca ccacttcaag aactctgtag	7320
caccgcctac atacctCgct ctgctaatacc tgttaccagt ggctgctgcc agtggcgata	7380
agtcgtgtct taccggggtt gactcaagac gatagttacc ggataaggcg cagcggtcgg	7440
gctgaacggg gggttcgtgc acacagccca gcttggagcg aacgacctac accgaactga	7500
gatacctaca gcgtgagcta tgagaaagcg ccacgcttcc cgaagggaga aaggcggaca	7560
ggatatccgtt aagcggCagg gtcggaacag gagagcgcac gagggagctt ccagggggaa	7620
acgcctggta tctttatagt cctgtcgggt ttcgccacct ctgacttgag cgtcgatttt	7680
tgtgatgctc gtcagggggg cggagcctat ggaaaaacgc cagcaacgcg gcctttttac	7740
ggttcctggc cttttgCtgg ccttttgctc acatggctcg acagatct	7788

<210> 141
 <211> 23
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> cebador SNE-S1

<400> 141	
ggttgggatt atccaaatg tga	23

<210> 142
 <211> 24
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> cebador SNE-AS1

<400> 142	
gcatcatcag aaagaatcat catg	24

<210> 143
 <211> 21
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> cebador SAR1-s

<400> 143	
cctctcttgt tctgctgc a	21

<210> 144
 <211> 21
 <212> ADN
 <213> Secuencia artificial

<220>

<223> cebador SAR1-AS

<400> 144
tatagtgagc cgccacacat g 21

<210> 145
<211> 45
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador PCR

<400> 145
ataggatcca ccatgtttat ttcttatta ttcttactc tcaact 45

<210> 146
<211> 37
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador PCR

<400> 146
atactcgagt tatgtgtaat gtaatttgac acccttg 37

<210> 147
<211> 45
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador PCR

<400> 147
ataggatcca ccatgtttat ttcttatta ttcttactc tcaact 45

<210> 148
<211> 36
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador PCR

<400> 148
acctccggat ttaatatatt gctcatattt tcccaa 36

<210> 149
<211> 13
<212> PRT
<213> Secuencia artificial

<220>
<223> extremo N-terminal de la proteína S del SRAS-CoV (aminoácidos 1 a 13)

<400> 149
Met Phe Ile Phe Leu Leu Phe Leu Thr Leu Thr Ser Gly
1 5 10

<210> 150
<211> 10
<212> PRT
<213> Secuencia artificial

<220>

<223> oligopéptido

<400> 150

Ser Gly Asp Tyr Lys Asp Asp Asp Asp Lys
1 5 10

<210> 151

<211> 34

<212> ADN

<213> Secuencia artificial

<220>

<223> cebador PCR

<400> 151

actagctagc ggatccacca tggtcatctt cctg 34

<210> 152

<211> 33

<212> ADN

<213> Secuencia artificial

<220>

<223> cebador PCR

<400> 152

agtatccgga ctgatgtac tgctcgctact tgc 33

<210> 153

<211> 59

<212> ADN

<213> Secuencia artificial

<220>

<223> oligonucléotide

<400> 153

tatgagcttt tttttttt ttttttggc atataaatag actcggcgcg ccatctgca 59

<210> 154

<211> 53

<212> ADN

<213> Secuencia artificial

<220>

<223> oligonucléotide

<400> 154

gatggcgcgc cgagtctatt tatatgcaa aaaaaaaaaa aaaaaaaagc tca 53

<210> 155

<211> 45

<212> ADN

<213> Secuencia artificial

<220>

<223> cebador PCR

<400> 155

atacgtagc ccatgtttat ttcttatta ttcttactc tca 45

<210> 156

<211> 40

<212> ADN
<213> Secuencia artificial

<220>
<223> cebador PCR

<400> 156
atagcgcgct cattatgtgt aatgtaattt gacacccttg 40

<210> 157
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador PCR

<400> 157
ccatttcaac aatttgccg 20

<210> 158
<211> 45
<212> ADN
<213> Secuencia artificial

<220>
<223> cebador PCR

<400> 158
ataggatccg cgcgctcatt atttatcgtc gtcattctta taatc 45

REIVINDICACIONES

1. Polipéptido aislado y purificado, caracterizado porque se trata del ectodominio de la proteína S de secuencia SEC ID nº 3, y porque está constituido por los aminoácidos que corresponden a las posiciones 1 a 1193 de la secuencia de aminoácidos de dicha proteína S o de los aminoácidos que corresponden a las posiciones 14 a 1193 de la secuencia de aminoácidos de dicha proteína S.
2. Ácido nucleico que codifica un polipéptido, según la reivindicación 1.
3. Vector de expresión recombinante, caracterizado porque comprende un ácido nucleico, según la reivindicación 2.
4. Vector de expresión recombinante, según la reivindicación 3, que codifica el ectodominio de la proteína S de secuencia SEC ID nº 3, caracterizado porque se selecciona entre los vectores contenidos en las capas bacterianas siguientes, depositadas en la Collection Nationale de Cultures de Microorganismes (CNCM), 25 rue du Docteur Roux, 75724 Paris Cedex 15
 - a) cepa nº I-3324, depositada el 22 de noviembre de 2004,
 - b) cepa nº 1-3327, depositada el 1 de diciembre de 2004,
 - c) cepa nº 1-3332, depositada el 1 de diciembre de 2004,
 - d) cepa nº I-3335, depositada el 1 de diciembre de 2004,
 - e) cepa nº I-3337, depositada el 1 de diciembre de 2004,
 - f) cepa nº 1-3339, depositada el 2 de diciembre de 2004, y
 - g) cepa nº I-3341, depositada el 2 de diciembre de 2004
5. Vector de expresión, según la reivindicación 3, caracterizado porque se trata de un vector viral, en forma de partícula viral o en forma de genoma recombinante.
6. Vector viral, en forma de partícula viral o en forma de genoma recombinante, según la reivindicación 5, caracterizado porque se trata de una partícula viral recombinante o de un genoma viral recombinante susceptible de ser obtenido por transfección de un plásmido según los apartados b) o d) a g) de la reivindicación 4, en un sistema celular apropiado.
7. Vector de expresión recombinante, según la reivindicación 3, caracterizado porque se trata de un vector lentiviral.
8. Vector de expresión recombinante, según la reivindicación 3, caracterizado porque se trata de un virus de la rubeola.
9. Vector de expresión recombinante, según la reivindicación 3, caracterizado porque se trata de un virus de la vacuna antivariólica.
10. Uso de un vector contenido en una cepa bacteriana, según la reivindicación 4, para la producción *in vitro*, en un sistema eucariota, del ectodominio de la proteína S del coronavirus asociado al SRAS, según la reivindicación 1.
11. Método de producción del ectodominio de la proteína S, según la reivindicación 1, en un sistema eucariota, que comprende una etapa de transfección de células eucariotas en cultivo por un vector seleccionado entre los vectores contenidos en las cepas bacterianas mencionadas en la reivindicación 5.
12. Célula eucariota aislada, genéticamente modificada, que expresa un polipéptido, según la reivindicación 1.
13. Célula, según la reivindicación 12, susceptible de ser obtenida mediante transfección por cualquiera de los vectores mencionados en los apartados d) o e) de la reivindicación 4.
14. Célula, según la reivindicación 13, caracterizada porque se trata de la célula FRhK4-Ssol-30, depositada en la CNCM el 22 de noviembre de 2004, bajo el nº I-3325.
15. Uso de un polipéptido, según la reivindicación 1, para detectar una infección por un coronavirus asociado al SRAS, a partir de una muestra biológica.
16. Método de detección de una infección por un coronavirus asociado al SRAS o de anticuerpos dirigidos contra un coronavirus asociado al SRAS, a partir de una muestra biológica, caracterizado porque la detección se efectúa por

ELISA utilizando un polipéptido, según la reivindicación 1, expresado en un sistema eucariota.

17. Método de detección, según la reivindicación 16, que comprende además una etapa de detección por ELISA que utiliza la proteína N recombinante.

18. Método, según la reivindicación 16 ó 17, caracterizado porque se trata de un método por ELISA doble epítipo, y porque el suero a ensayar se mezcla con el antígeno de revelado, siendo puesta después dicha mezcla en contacto con el antígeno fijado sobre un soporte sólido.

19. Complejo inmune aislado formado por un polipéptido, según la reivindicación 1, y un anticuerpo dirigido específicamente contra un epítipo del coronavirus asociado al SRAS.

20. Kit o estuche de detección de un coronavirus asociado al SRAS, caracterizado porque comprende el menos un agente reactivo seleccionado del grupo constituido por: un polipéptido, según la reivindicación 1, un ácido nucleico, según la reivindicación 2, una célula, según cualquiera de las reivindicaciones 12 a 14.

21. Composición inmunógena y/o de vacuna, caracterizada porque comprende un polipéptido recombinante, según la reivindicación 1, obtenido en un sistema de expresión eucariota.

22. Composición inmunógena y/o de vacuna, caracterizada porque comprende un vector o un virus recombinante, según cualquiera de las reivindicaciones 3, 4, y 5 a 9.

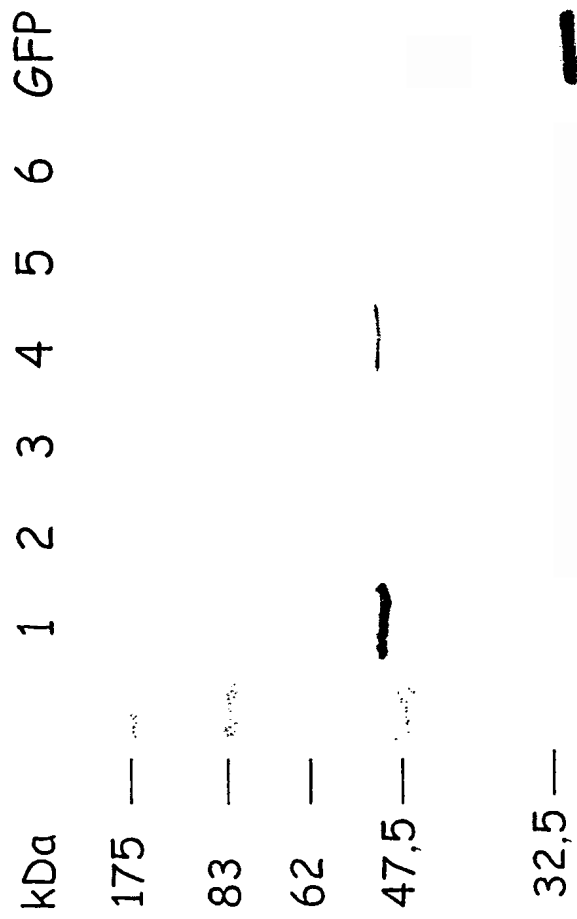


FIGURA 1

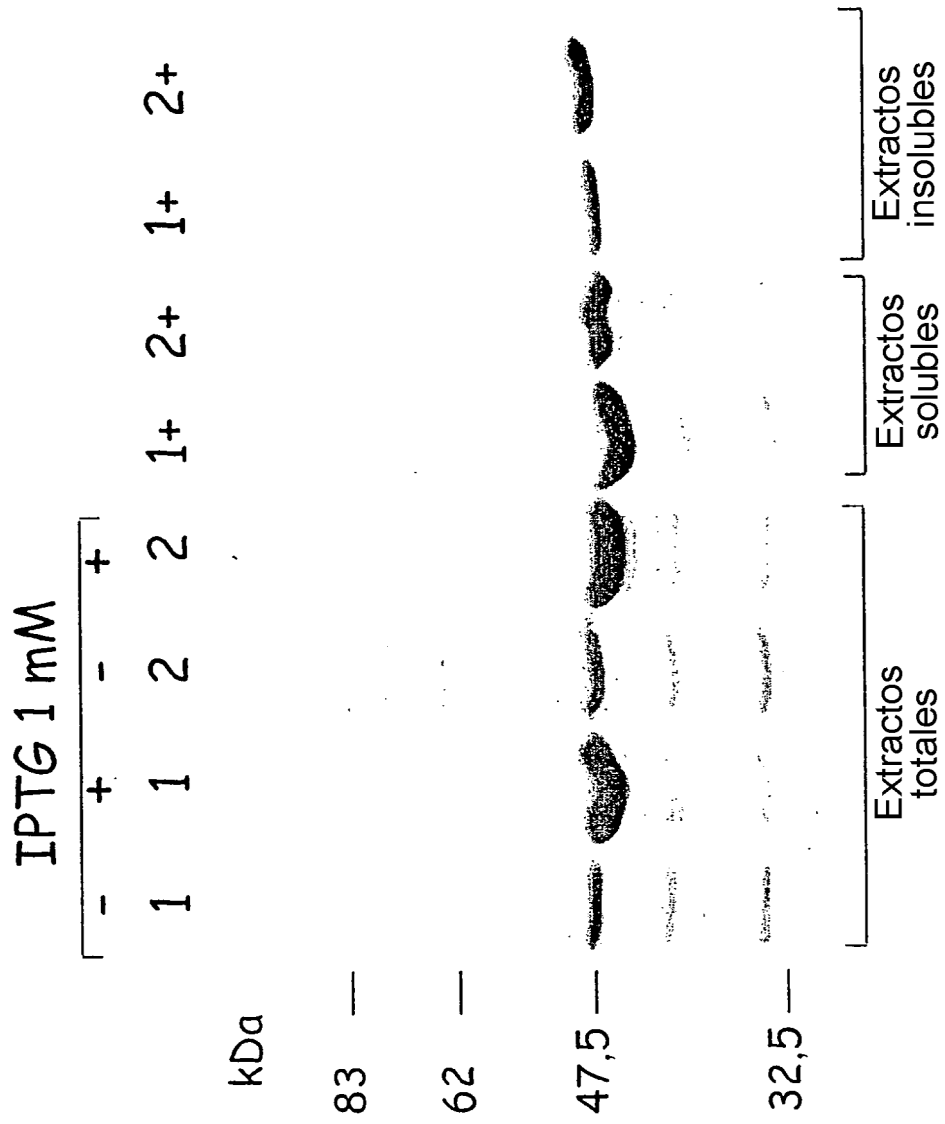


FIGURA 2



FIGURA 3

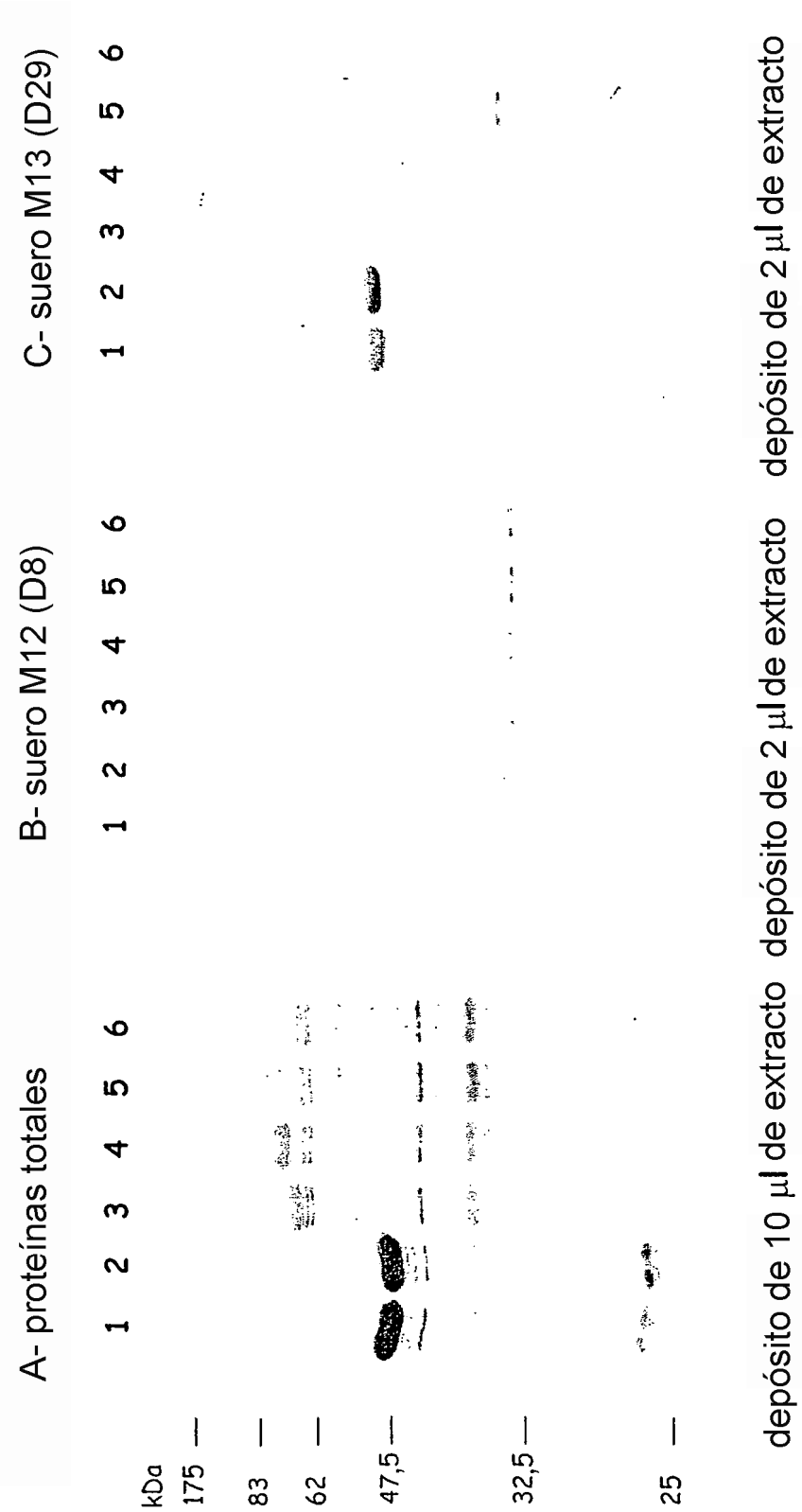


FIGURA 4

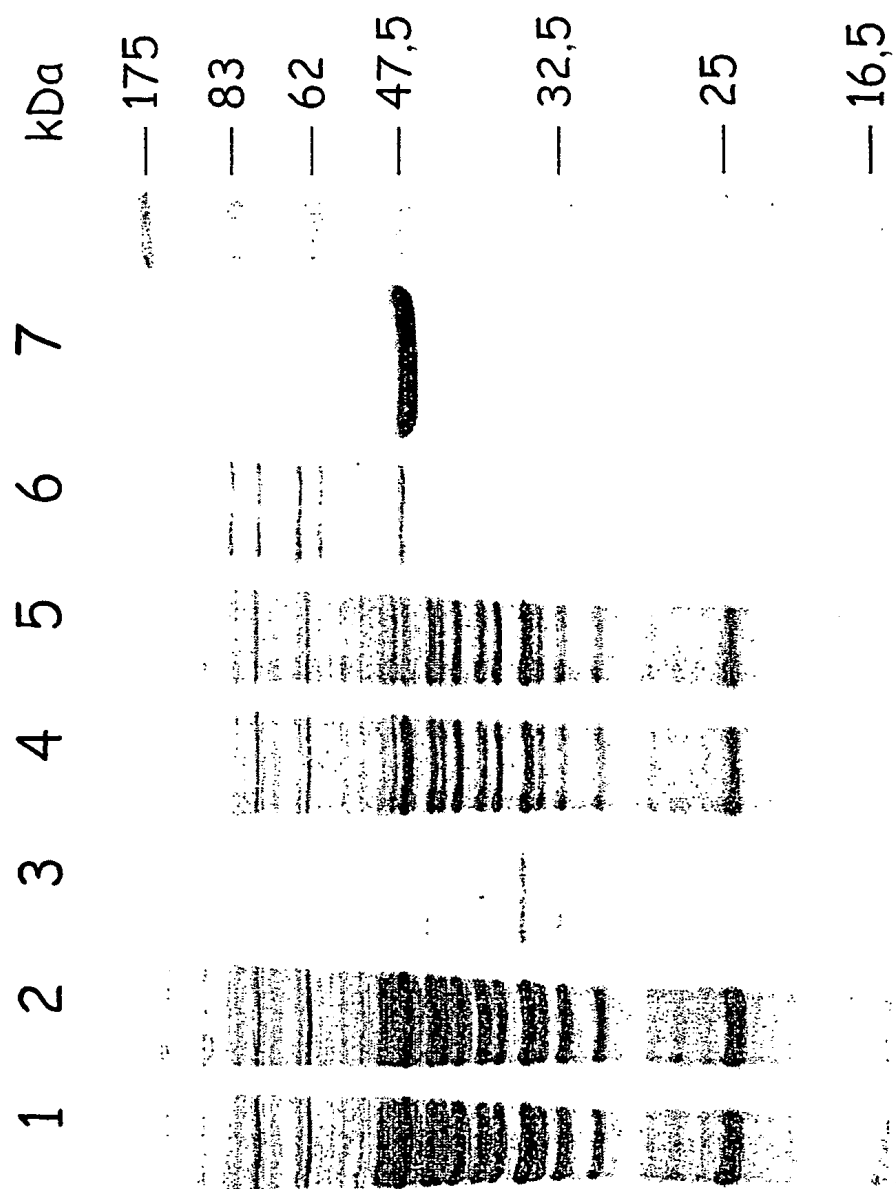


FIGURA 5

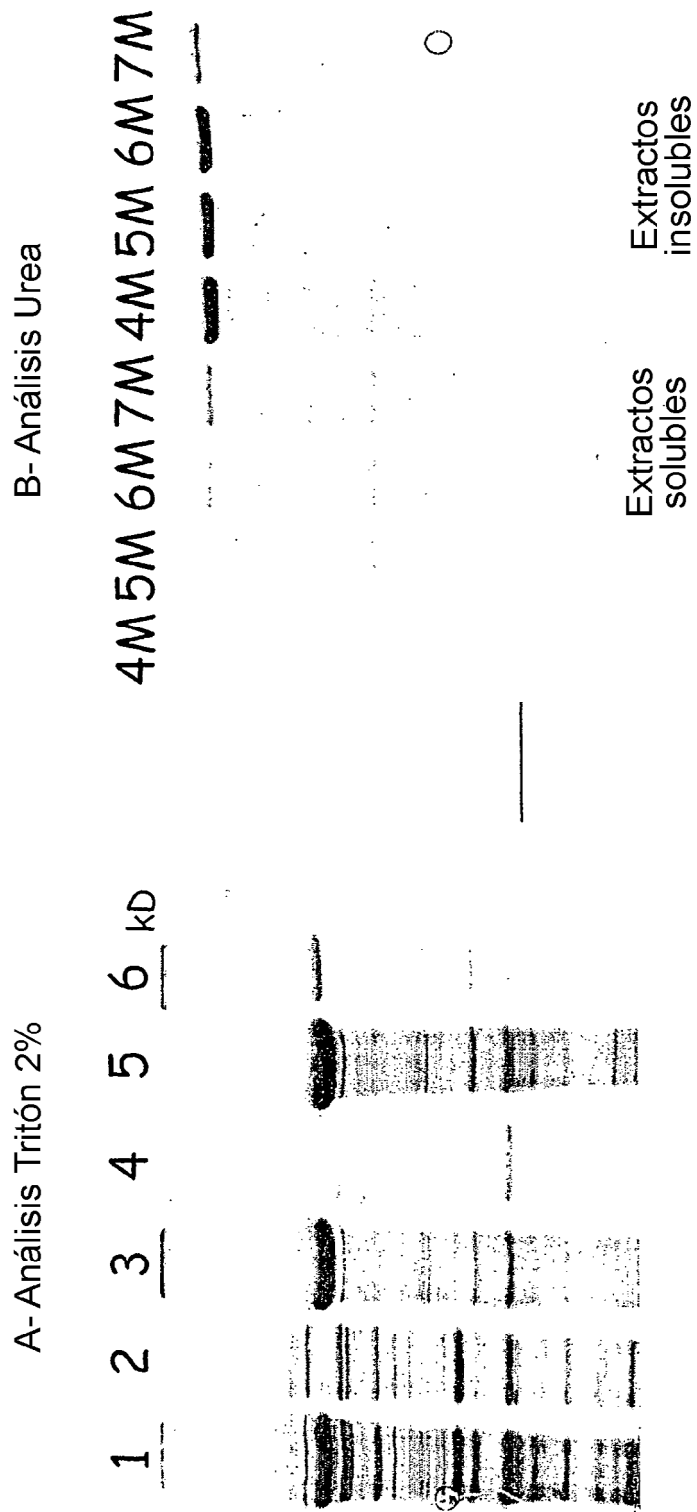


FIGURA 6

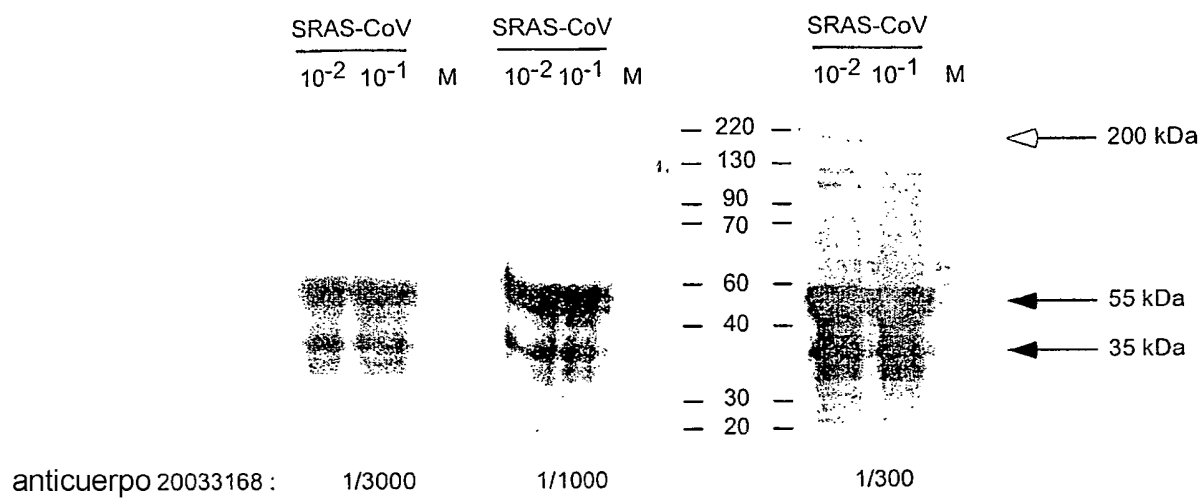


FIGURA 7

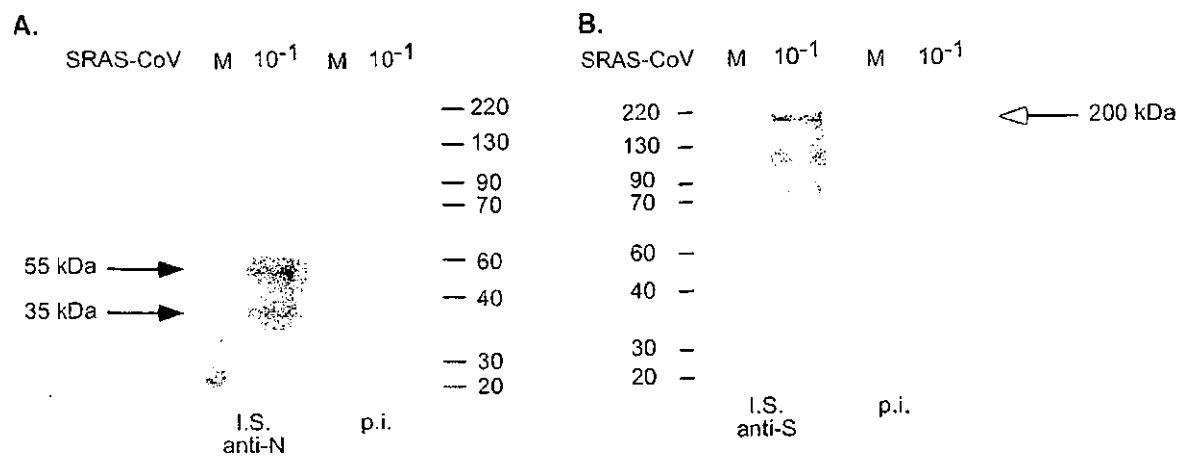
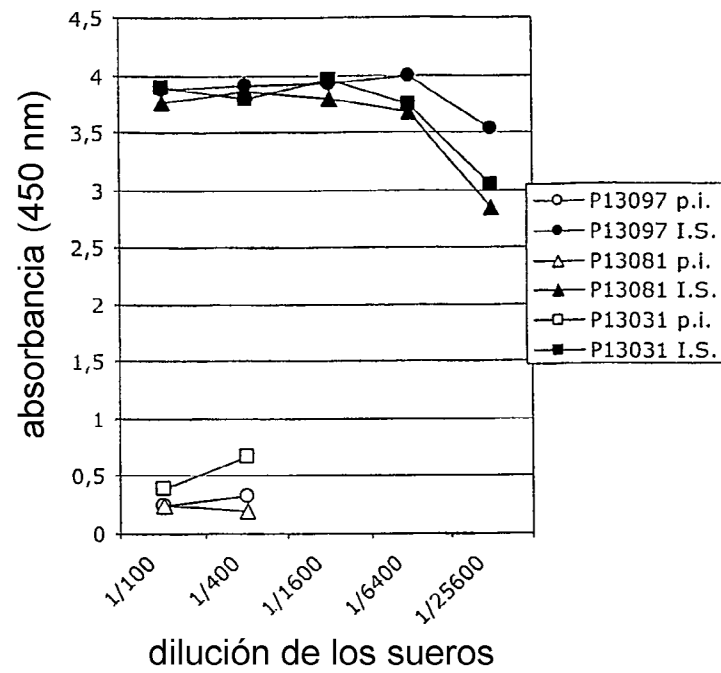


FIGURA 8

A



B

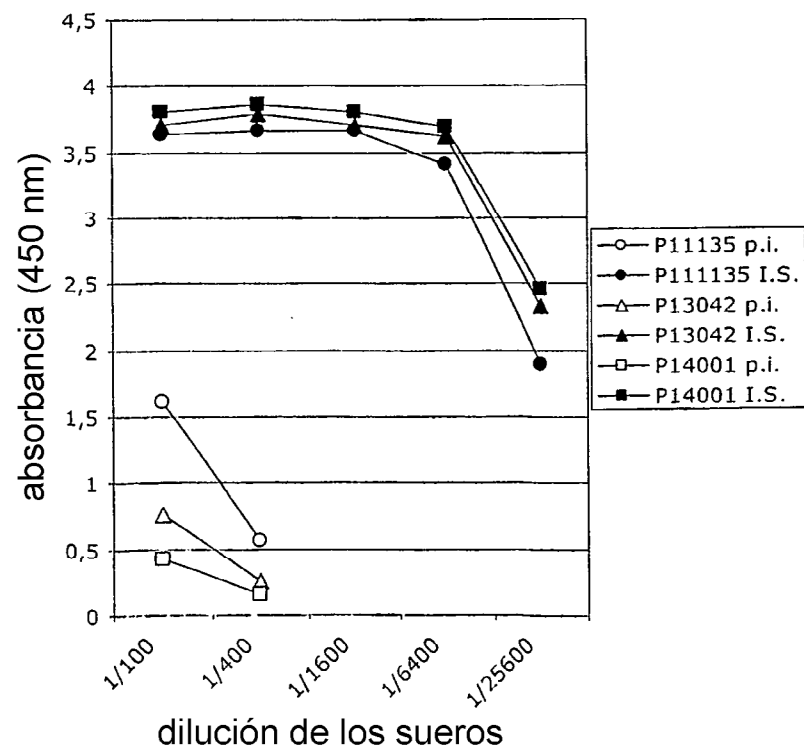


FIGURA 9

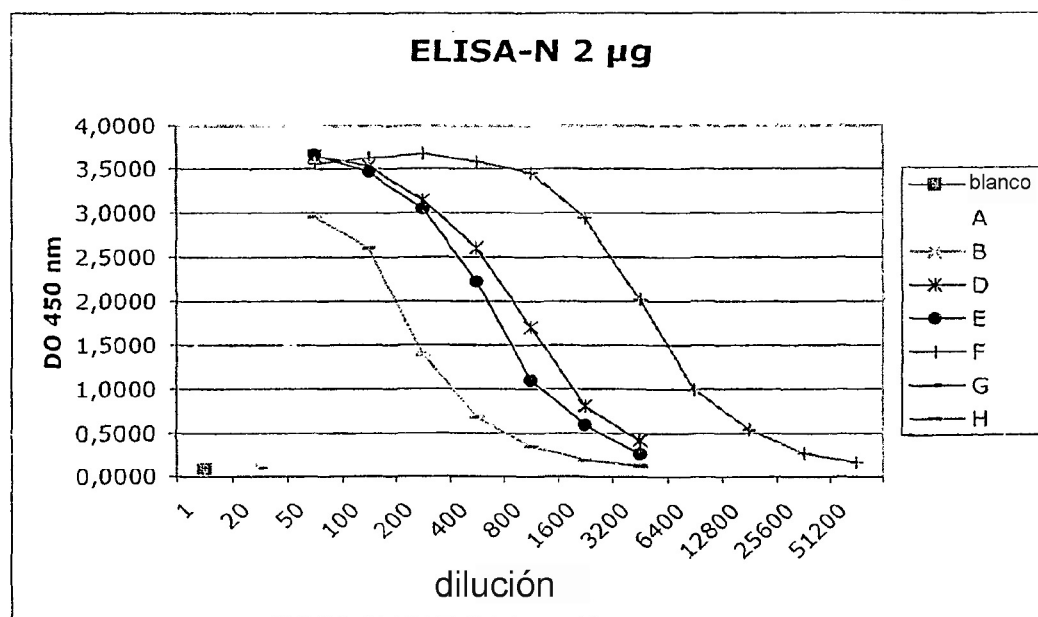
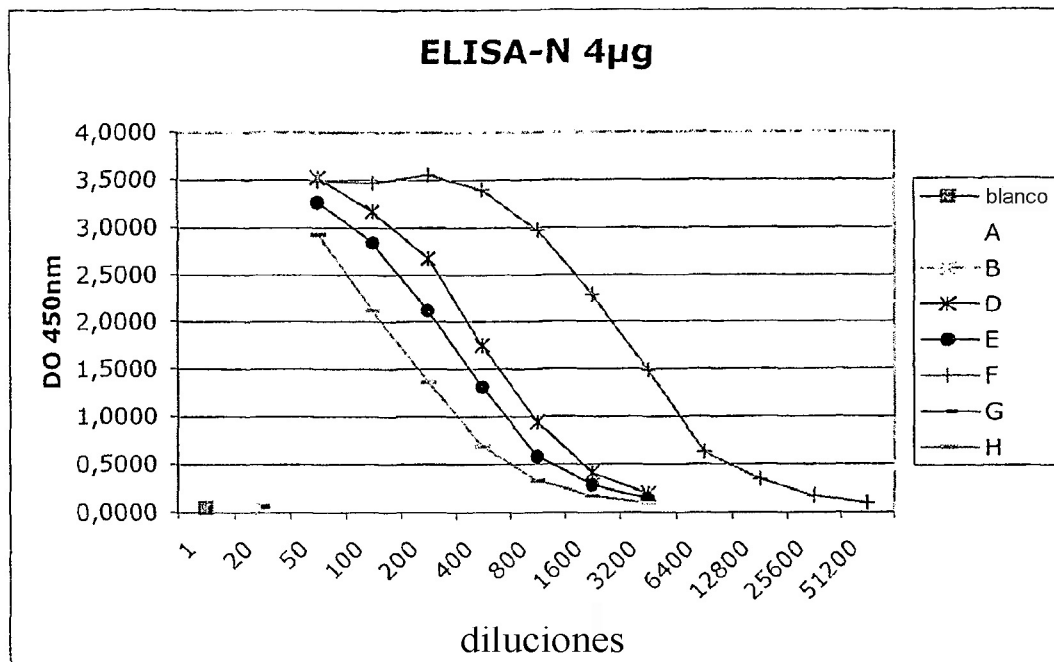


FIGURA 10a

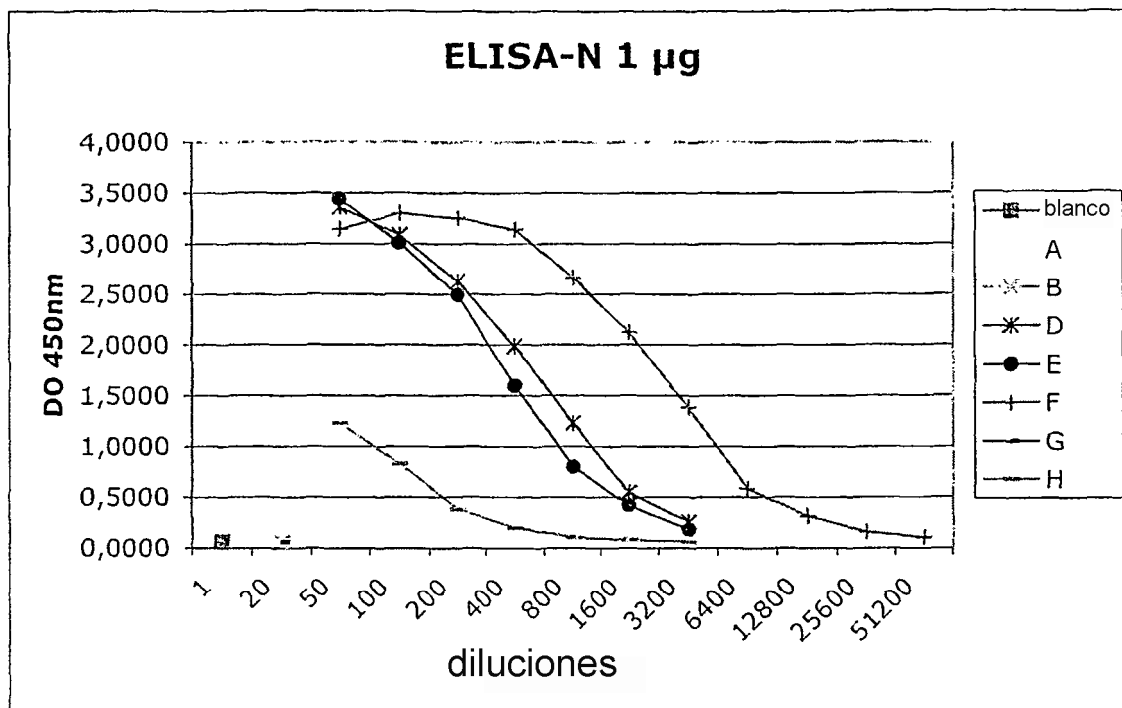


FIGURA 10b

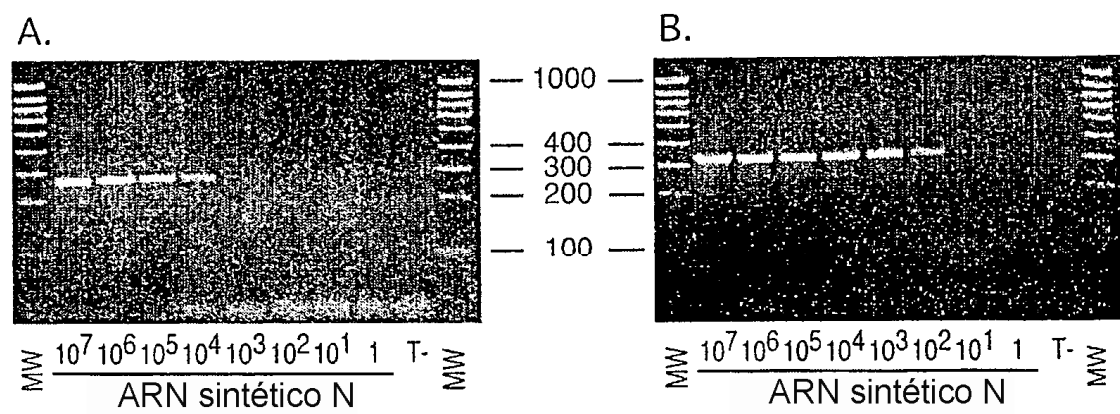


FIGURA 11

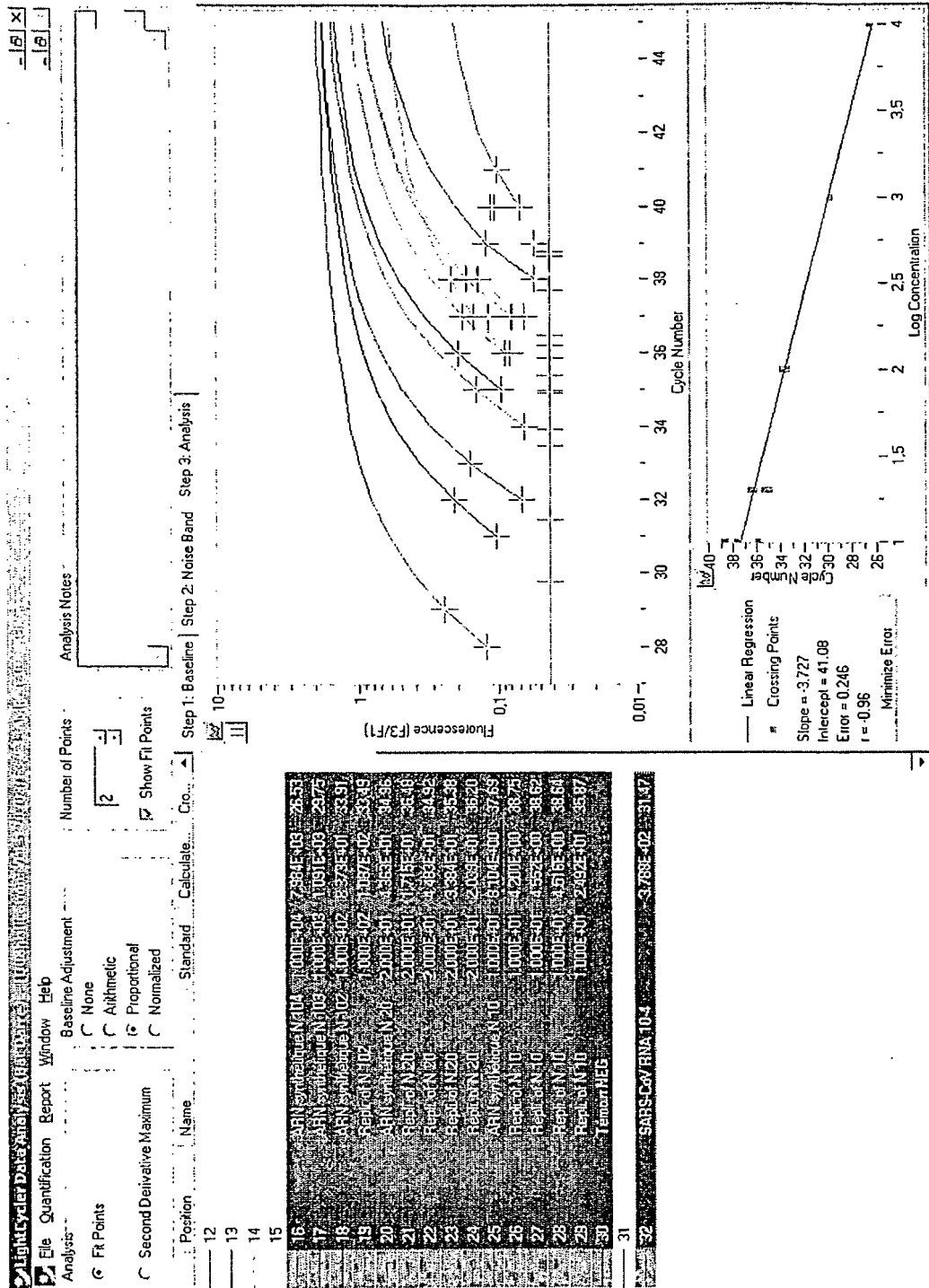


FIGURA 12

ES 2 396 127 T3

```

                >< ScrFI                                >< XhoII
                >< MvaI                                > < TthHB8I >< NdeII
>< EcoRII      > < TaqI                                >< MflI
                >< Ecl136I                            >< Sau3AI >< MboI
>< DsaV        >< NdeII                                >< DpnII
                >< BstOI                                >< MboI>< MnlI>< DpnI
                >< BstNI                                >< DpnII >< BstYI
                >< BsiLI                                >< DpnI >< BspAI
>< BsaJI        >< BspAI                                >< Bsp143I
                >< ApyI                                >< Bsp143I>< BglII
ATATTAGGTT TTTACCTACC CAGGAAAAGC CAACCAACCT CGATCTCTTG TAGATCTGTT CTCTAAACGA
    10          20          30          40          50          60          70

                >< VneI
                >< SphI
                >< SnaI
                >< RmaI
                >< PaeI >< SduI
                >< NspI >< NspII
                >< NspHI >< HgiAI
                >< NlaIII >< Bsp1286I
                >< MaeI >< BmyI
>< Tru9I
>< MseI        >< BbvI                                >< ApaLI
                >< DraI >< AluI > < Fnu4HI >< Alw44I
                >< DraI >< AluI >< Alw21I
ACTTTAAAT CTGTGTAGCT GTCGCTCGGC TGCATGCCTA GTGCACCTAC GCAGTATAAA CAATAATAAA
    80          90          100          110          120          130          140

                >< SfcI
                >< PstI
                >< MnlI
                >< Ksp632I
                >< HindII > < MboII >< EarI
                >< HincII >< MaeIII >< Eam1104I
TTTACTGTC GTTGACAAGA AACGAGTAAC TCGTCCCTCT TCTGCAGACT GCTTACGGTT TCGTCCGTGT
    150          160          170          180          190          200          210

                >< TthHB8I >< StyI
>< TaqI        >< RmaI >< ScrFI
>< Sau3AI      >< MaeI >< NciI
>< NdeII       >< Eco14I >< MspI
>< MboI        >< Eco130I >< MaeIII
>< DpnII       >< BssTII >< HpaII
>< DpnI        >< BsaJI >< HapII
>< BspAI       >< BlnI >< DsaV
>< Bsp143I     >< AvrII >< BcnI
TGCAGTCGAT CATCAGCAT CCTAGGTTTC GTCCGGGTGT GACCGAAAGG TAAGATGGAG AGCCTTGTTT
    220          230          240          250          260          270          280

                >< RmaI
                >< Esp3I >< MaeII
>< HindII      >< MaeII> < Eco57I >< BsmAI >< MaeI
>< HincII > < AflIII > < DdeI >< Alw26I >< BsmBI
TTGGTGTCAA CGAGAAAACA CACGTCCAAC TCAGTTTGCC TGTCCCTCAG GTTAGAGACG TGCTAGTGCG
    290          300          310          320          330          340          350

```

FIGURA 13.1

```

>< Sau96I
  >< PssI
    >< PstI
      >< NspIV
        >< MnlI
          >< HaeIII
            >< EcoO109I
              >< DraII>< MboII >< PmlI
                >< MnlI >< Cfr13I >< PmaCI
                  >< Ksp632I >< BsuRI > < MaeII
                    >< HinfI >< BsiZI>< EcoNI >< Eco72I
                      >< EarI >< BshI >< BslI >< BsaAI
                        >< PleI >< Eam1104I>< AsuI >< BsiYI>< BbrPI >< MnlI
TGGCTTCGGG GACTCTGTGG AAGAGGCCCT ATCGGAGGCA CGTGAACACC TCAAAAATGG CACTTGTGGT
  360      370      380      390      400      410      420

                                >< Tru9I
                                >< SfaNI
                                >< MseI
>< RmaI >< RsaI >< Csp6I >< BspWI >< MaeII
>< MaeI >< AluI >< AfaI >< AluI >< MaeII
CTAGTAGAGC TGGAAAAAGG CGTACTGCCC CAGCTTGAAC AGCCCTATGT GTTCATTAAA CGTTCTGATG
  430      440      450      460      470      480      490

                                >< RsaI
                                >< MseI
                                >< Csp6I
                                >< BsmI BsiEI ><
>< Tru9I >< GdiII >< RsaI
>< MseI >< EaeI >< Csp6I
>< Esp4I >< BsuRI >< BsmI BsiEI ><
>< AflIII >< BshI >< AluI >< BscCI >< AfaI
CCTTAAGCAC CAATCACGGC CACAAGGTCG TTGAGCTGGT TGCAGAAATG GACGGCATTC AGTACGGTCG
  500      510      520      530      540      550      560

                                >< NspI
                                >< NspHI
                                >< NlaIII
                                >< BslI
                                >< BsiYI
                                >< MboII
                                >< MboII
>< AciI >< AfaI >< AflIII >< MunI >< AciI
TAGCGGTATA ACGTGGGAG TACTCGTGCC ACATGTGGGC GAAACCCCAA TTGCATACCG CAATGTTCTT
  570      580      590      600      610      620      630

                                >< TthHB8I
                                >< TaqI
                                >< Sau3AI
                                >< NdeII
                                >< MboI
                                >< DpnII
                                >< DpnI
                                >< ClaI
                                >< Bsu15I
                                >< BspDI
                                >< BspAI
                                >< Bsp143I
                                >< Bsp106I
                                >< BsiXI
                                >< BscI>< SfaNI DdeI ><
                                >< BscBI >< AluI >< BanIII BfrI ><
CTTCGTAAGA ACGGTAATAA GGGAGCCGGT GGTGATAGCT ATGGCATCGA TCTAAAGTCT TATGACTTAG
  640      650      660      670      680      690      700

```

FIGURA 13. 2

```

    >< Sau3AI
    >< NdeII
    >< MboI
    >< HphI
    >< DpnII
    >< BspAI
    >< AlwI>< DpnI
    >< AluI
    >< Bsp143I
    >< MboII
    >< BsrI
    >< DdeI
    >< VneI ><
    >< SnoI ><
    >< NlaIII
    >< ApaLI ><
    >< Alw44I ><
    GTGACGAGCT TGGCACTGAT CCCATTGAAG ATTATGAACA AAAGTGAAC ACTAAGCATG GCAGTGGTGC
    710 720 730 740 750 760 770

    >< SstI
    >< SduI
    >< SacI
    >< NspII
    >< MnlI
    >< HgiAI
    >< Eco24I
    >< TthHB8I
    >< PalI ><
    >< NspII
    >< Ecl136II
    >< SalI
    >< HaeIII ><
    >< HgiAI
    >< Bsp1286I
    >< RtrI
    >< Cfr13I ><
    >< DraIII
    >< BmyI
    >< HindII
    >< BsuRI ><
    >< Bsp1286I
    >< BanII
    >< HincII
    >< BsiZI ><
    >< BmyI
    >< Alw21I
    >< BsgI
    >< BshI ><
    >< Alw21I
    >< AluI
    >< MaeIII
    >< AccI
    >< AsuI ><
    ACTCCGTGAA CTCACCTCGTG AGCTCAATGG AGGTGCAGTC ACTCGCTATG TCGACAACAA TTTCTGTGGC
    780 790 800 810 820 830 840

    >< Thai
    >< Thai
    >< MvnI
    >< MvnI
    >< HinPII
    >< Hin6I
    >< HhaI
    >< VneI
    >< CfoI
    >< SnoI
    >< SduI
    >< Eco64I
    >< Csp6I
    >< BstUI
    >< BstUI
    >< NspII ><
    >< BscBI
    >< BstUI
    >< HgiAI ><
    >< BanI
    >< Bsp50I
    >< Bsp1286I ><
    >< Asp718
    >< Bsp50I
    >< BmyI
    >< AfaI
    >< AciI
    >< ApaLI
    >< AccB1I
    >< AccII
    >< Alw44I
    >< Acc65I
    >< MnlI
    >< SfaNI
    >< AccII
    >< Alw21I ><
    CCAGATGGGT ACCCTCTTGA TTGCATCAAA GATTTTCTCG CACGCGCGGG CAAGTCAATG TGCACTCTTT
    850 860 870 880 890 900 910

    >< TthHB8I
    >< TthHB8I
    >< TaqI
    >< TaqI
    >< MnlI
    >< Ksp632I
    >< NlaIII ><
    >< HinfI>< PleI
    >< NlaIII
    >< Eam1104I
    >< MboII
    >< MaeIII
    >< EcoRII ><
    >< EarI
    >< BbvI>< AccI >< Fnu4HI
    >< DsaV ><
    CCGAACAACCT TGATTACATC GAGTCGAAGA GAGGTGTCTA CTGCTGCCGT GACCATGAGC ATGAAATTGC
    920 930 940 950 960 970 980

    >< TthHB8I
    >< TaqI
    >< SfuI
    >< NspV>< Tru9I
    >< LspI>< MseI
    >< ScrFI
    >< HinPII

```

FIGURA 13. 3

ES 2 396 127 T3

```

>< MvaI      >< Hin6I      >< SduI      >< Csp45I
>< Ecl136I   >< HhaI      >< NspII     >< BstBI
>< BstOI     >< HaeII     >< HgiAI     >< Bsp119I
>< BstNI     >< Eco47III   >< Bsp1286I  >< BsiCI
>< BsiLI     >< CfoI      >< BmyI      >< Bpu14I
>< ApyI >< DdeI >< Bsp143II >< AluI      >< Alw21I    >< AsuII
CTGGTTCAC T GAGCGCTCTG ATAAGAGCTA CGAGCACCAG ACACCCTTCG AAATTAAGAG TGCCAAGAAA
      990      1000      1010      1020      1030      1040      1050

                                >< Tru9I
                                >< MseI
                                >< BsmI
                                >< BscCI
                                >< MnlI
TTTGACACTT TCAAAGGGGA ATGCCCAAAG TTTGTGTTTC CTCTTAAC TC AAAAGTCAAA GTCATTCAAC
      1060      1070      1080      1090      1100      1110      1120

>< PmlI
>< PmaCI
>< MaeII
>< Eco72I
>< BsaAI      >< NlaIII      >< RsaI
>< BbrPI      >< Bst1107I >< Csp6I
>< AflIII     >< MnlI>< DdeI      >< AccI      >< AfaI
CACGTGTGTA AAAGAAAAAG ACTGAGGGTT TCATGGGGCG TATACGCTCT GTGTACCCTG TTGCATCTCC
      1130      1140      1150      1160      1170      1180      1190

>< SfaNI
>< MaeIII     >< AccI      NlaIII ><
ACAGGAGTGT AACAAATATGC ACTTGCTAC CTTGATGAAA TGTAATCATT GCGATGAAGT TTCATGGCAG
      1200      1210      1220      1230      1240      1250      1260

                                >< SinI
                                >< Sau96I
                                PssI ><
                                >< Psp5II
                                >< PpuMI
                                >< NspIV
                                >< NspHII
                                >< Eco47I
                                >< DraII
                                >< Cfr13I
                                >< BsiZI
                                >< Bme18I
                                >< AvaII
                                >< AsuI
>< MaeII     EcoO109I ><AflIII >
ACGTGCGACT TTCTGAAAGC CACTTGTGAA CATTGTGGCA CTGAAAATTT AGTTATTGAA GGACCTACTA
      1270      1280      1290      1300      1310      1320      1330

                                Van91I ><
                                SinI ><
                                Sau96I ><
                                PflMI ><
                                NspIV ><
                                NspHII >
                                Eco47I ><
                                Cfr13I ><
                                BsiI ><
                                BsiZI ><
                                BsiYI ><
                                Bme18I ><
                                AvaII ><
                                AsuI ><

>< RsaI
>< NspI
>< NlaIV
>< NlaIII
>< NspHI>< KpnI
>< Eco64I
>< Csp6I
>< BscBI
>< BanI
>< Asp718
>< AfaI
>< AccBII

```

FIGURA 13. 4

```

    >< Acc65I          >< SfcI          >< NlaIII          AccB7I ><
CATGTGGGTA CCTACCTACT AATGCTGTAG TGAAAATGCC ATGTCCTGCC TGTCAAGACC CAGAGATTGG
    1340          1350          1360          1370          1380          1390          1400

                                >< TthHB8I
                                >< TaqI>< MnlI
                                >< HinfI
    >< DdeI
ACCTGAGCAT ACTGTTGCAG ATTATCACAA CCACTCAAAC ATTGAAACTC GACTCCGCAA GGGAGGTAGG
    1410          1420          1430          1440          1450          1460          1470

    >< RmaI          NlaIV ><
    >< MnlI          >< BsrI
    >< MaeI          >< BbvI          >< Fnu4HI          BscBI ><
ACTAGATGTT TTGGAGGCTG TGTGTTTGCC TATGTTGGCT GCTATAATAA GCGTGCCTAC TGGGTTCCCTC
    1480          1490          1500          1510          1520          1530          1540

                                XhoII ><
                                Sau3AI ><
                                NdeII ><
                                MflI ><
                                MboI ><
                                >< MaeIII          >< Eco3II          DpnII ><
                                >< Pali          >< HaeIII          >< BsrI          >< MnlI DpnI >
    >< RmaI          >< BsuRI          >< BsrI          >< BsmAI          BstYI ><
    >< MnlI          >< DdeI          >< BspWI          >< BsaI>< HphI          BspAI ><
    >< MaeI          >< BshI>< BglI          >< Alw26I          Bsp143I >
GTGCTAGTGC TGATATTGGC TCAGGCCATA CTGGCATTAC TGGTGACAAAT GTGGAGACCT TGAATGAGGA
    1550          1560          1570          1580          1590          1600          1610

                                >< Tru9I
                                >< MseI
                                >< MaeII          >< Tru9I
                                >< HpaI          >< MnlI
                                >< HindII          >< Ksp632I
                                >< HinfI >< PleI >< HincII          >< EarI
    >< AlwI >< DdeI          >< AflIII          >< MseI          >< Eam1104I
TCTCCTTGAG ATACTGAGTC GTGAACGTGT TAACATTAAC ATTGTTGGCG ATTTTCATTT GAATGAAGAG
    1620          1630          1640          1650          1660          1670          1680

    >< MboII          PleI ><
    >< BstXI          >< SfaNI          >< HinfI
GTTGCCATCA TTTTGGCATC TTTCTCTGCT TCTACAAGTG CCTTTATTGA CACTATAAAG AGTCTTGATT
    1690          1700          1710          1720          1730          1740          1750

                                >< StyI
                                >< MaeIII
                                >< EcoT14I
                                >< PleI          >< Eco130I
                                >< MaeIII          >< BssTII          BslI ><
                                >< HinfI>< AciI          >< BsaJI          BsiYI ><
ACAAGTCTTT CAAAACCATT GTTGAGTCCT GCGGTAACTA TAAAGTTACC AAGGGAAAGC CCGTAAAAGG
    1760          1770          1780          1790          1800          1810          1820

    >< Sau3AI          >< Van9II
    >< NdeII          >< PflMI
    >< MboI          >< DraIII
    >< DpnII          >< BslI
    >< DpnI >< Tru9I          >< BsiYI
    >< BspAI >< MseI          >< BbvI          >< MnlI
    >< Bsp143I          >< AccB7I          Fnu4HI ><

```

FIGURA 13. 5

```

TGCTTGGAAC ATTGGACAAC AGAGATCAGT TTTAACACCA CTGTGTGGTT TTCCCTCACA GGCTGCTGGT
1830      1840      1850      1860      1870      1880      1890

                >> Thai
                >> SfaNI
                >> MvnI
                >> HinPII
                >< HinPII
                >< Hin6I
                >< Hin6I
                    >< HhaI
>< Sau3AI      >< HhaI
>< NdeII      >< CfoI
>< MboI      >< CfoI
>< DpnII      >< BstUI
    >< DpnI    >< BssHII
>< BspAI      >< Bsp50I
    >< Bsp143I >< AccII
GTTATCAGAT CAATTTTTCG GCGCACACTT GATGCAGCAA ACCACTCAAT TCCTGATTG CAAAGAGCAG
1900      1910      1920      1930      1940      1950      1960

                >> TthHB8I
                >> StyI
                >> NcoI
                >< HindII
                >< HincII
                >< HinII
                >< EcoT14I
                >< Eco57I
                >< TaqI>< Eco130I
                >< SalI >< DsaI
                >< RtrI >< BssTII
                >< BsaHI
                >< BbiII>< NlaIII
                >< AcyI >< HgaI

>< MaeIII
    >< BbvI
CTGTCACCAT ACTTGATGGT ATTTCTGAAC AGTCATTACG TCTTGTGCGAC GCCATGGTTT ATACTTCAGA
1970      1980      1990      2000      2010      2020      2030

                >> RsaI
                >< Csp6I
    >< BspMI
                >< MaeIII >< BsrI >< AfaI >< DdeI
CCTGCTCACC AACAGTGTCA TTATTATGGC ATATGTAAC TGTGGTCTTG TACAACAGAC TTCTCAGTGG
2040      2050      2060      2070      2080      2090      2100

                >< StuI
                >< Pali
                >< MaeIII
                >< Ecol47I
    >< SduI
    >< NspII
    >< Bsp1286I
    >< BmyI
    >< DdeI
    >< BsuRI
    >< BshI
    >< AatI
    >< MnlI
    >< DdeI ><
    >< BfrI ><
TTGTCTAATC TTTTGGGCAC TACTGTTGAA AAACTCAGGC CTATCTTTGA ATGGATTGAG GCGAAACTTA
2110      2120      2130      2140      2150      2160      2170

                >< TfiI
                >< HinfI
                >< FokI
    >< SfaNI >< BsqI
    >< FokI
GTGCAGGAGT TGAATTTCTC AAGGATGCTT GGGAGATTCT CAAATTTCTC ATTACAGCTG TTTTTCACAT
2180      2190      2200      2210      2220      2230      2240

```

FIGURA 13. 6

```

Tru9I >>
MseI >>
HpaI >
HindII >
HincII >
>< Eco57I
CGTCAAGGGT CAAATACAGG TTGCTTCAGA TAACATCAAG GATTGTGTAA AATGCTTCAT TGATGTTGTT
2250      2260      2270      2280      2290      2300      2310

>< Sau3AI
>< NdeII
>< MboI
> < MaeIII
>< FbaI
>< DpnII
>< DpnI
>< BspAI
>< Bsp143I
>< TthHB8I
>< TaqI
AACAAGGCAC TCGAAATGTG CATTGATCAA GTCACATATCG CTGGCGCAAA GTTGGGATCA CTCAACTTAG
2320      2330      2340      2350      2360      2370      2380

>< Sau3AI
>< NdeII
>< DpnII
>< DpnIMboII ><
DdeI ><
>< HinP1I
>< Hin6I
>< HhaI
>< CfoI
>< BspAI
>< BbsI
>< PvuII
>< MaeII
>< Bst1107I
>< BsaAI
>< BbvI
>< Fnu4HI
>< Fnu4HI
>< HphI
>< DrdI
>< AccI
>< AluI
GTGAAGTCTT CATCGCTCAA AGCAAGGGAC TTTACCGTCA GTGTATACGT GGCAAGGAGC AGCTGCAACT
2390      2400      2410      2420      2430      2440      2450

>< Tru9I
>< NlaIV
>< MseI
>< MnlI
>< Esp4I
>< Eco64I
>< BscBI
>< NlaIII >< BanI
>< AflII
>< BbvI
>< AccBI
>< MaeIII
>< TfiI
>< HinfI
>< HphI
>< AfaI
ACTCATGCCT CTTAAGGCAC CAAAAGAAGT AACCTTTCTT GAAGGTGATT CACATGACAC AGTACTTACC
2460      2470      2480      2490      2500      2510      2520

> < XhoI
>< TthHB8I
>< TthHB8I>< TaqI
> < SlaI
> < PaeR7I
> < NspIII
>< HphI >< HinfI
> < Eco88I
> < CcrI
>< Esp3I >< BsaHI
> < BcoI
>< BsmAI >< BbiII
> < AvaI >< HgaI
>< TaqI > < Ama87I>< BsmBI
>< Alw26I >< AcyI
>< DdeI>< MnlI
TCTGAGGAGG TTGTTCTCAA GAACGGTGAA CTCGAAGCAC TCGAGACGCC CGTTGATAGC TTCACAAATG
2530      2540      2550      2560      2570      2580      2590

```

FIGURA 13. 7


```

                                >< PstI >< NlaIII
                                >< HaeIII >< MnlI
                                >< BsuRI >< DdeI >< Tru9I
                                >< BshI >< BfrI >< MseI
>< AluI >< BsrI
GAGCTATCGT TGGCACACCA GTCTGTGTAA ATGGCCTCAT GCTCTTAGAG ATTAAGGACA AAGAACAATA
2600 2610 2620 2630 2640 2650 2660

                                >< VneI
                                Tru9I ><
                                >< SnaI
                                >< SduI
                                >< NspII
                                MseI ><
                                >< HgiAI
                                Bsp1286I >< BslI ><
                                BsiYI ><
                                >< BmyI
                                >< ApaLI
                                >< Tru9I >< Alw44I
                                >< MseI >< Alw21I
CTGCGCATTTG TCTCCTGGTT TACTGGCTAC AAACAATGTC TTTCGCTTAA AAGGGGGTGC ACCAATTAAA
2670 2680 2690 2700 2710 2720 2730

                                >< TfiI
>< MaeIII >< MboII >< MaeIII >< HinfI AluI ><
GGTGTAACCT TTGGAGAAGA TACTGTTTGG GAAGTTCAAG GTTACAAGAA TGTGAGAATC ACATTTGAGC
2740 2750 2760 2770 2780 2790 2800

                                >< RsaI
                                >< NlaIV
                                MaeIII ><
                                >< MspI >< KpnI
                                >< HpaII
                                >< HapII
                                >< Eco64I
                                >< SduI >< Csp6I
                                >< NspII >< BscBI
                                >< HgiAI >< BanI
                                >< Bsp1286I >< Asp718
                                >< BmyI >< HinfI >< AfaI
                                >< Alw21I >< AccB1I
                                >< AccI >< Acc65I
TTGATGAACG TGTTGACAAA GTGCTTAATG AAAAGTGTCT TGTCTACACT GTTGAATCCG GTACCGAAGT
2810 2820 2830 2840 2850 2860 2870

                                >< Sau3AI
                                >< NdeII
                                >< MboI
                                >< DpnII
                                >< DpnI
                                >< NspI
                                >< NspHI
                                >< MboII >< BspAI
                                >< NlaIII >< BsrI >< Bsp143I
                                >< MnlI >< AlwNI >< BbsI >< AlwNI
>< DdeI >< MnlI >< AlwNI >< BbsI >< AlwNI
TACTGAGTTT GCATGTGTTG TAGCAGAGGC TGTGTGAAG ACTTTACAAC CAGTTTCTGA TCTCCTTACC
2880 2890 2900 2910 2920 2930 2940

                                >< Sau3AI
                                >< NdeII
                                >< MboI
                                >< DpnII
                                >< DpnI
                                >< BspAI

```

FIGURA 13. 8

```

    >< NlaIII>< Bsp143I          >< AluI          >< SfaNI
AACATGGGTA TTGATCTTGA TGAGTGGAGT GTAGCTACAT TCTACTTATT TGATGATGCT GGTGAAGAAA
    2950          2960          2970          2980          2990          3000          3010

                                >< SfaNI
                                >< MnlI
    >< MboII          >< GsuI          >< Ksp632I          >< MnlI
                                >< BsaAI          >< EarI          >< MboII
    >< HphI          >< MaeII>< BpmI          >< MnlI          >< Eam1104I          >< MboII
ACTTTTCATC ACGTATGTAT TGTTCCCTTT ACCCTCCAGA TGAGGAAGAA GAGGACGATG CAGAGTGTGA
    3020          3030          3040          3050          3060          3070          3080

                                >< RsaI
                                >< RsaI
    >< NlaIII
                                >< MnlI          >< FokI
                                >< Csp6I          Eco31I ><
                                >< Csp6I          >< MamI BsmAI ><
                                >< MboII          >< AfaI          >< BsiBI BsaI ><
    >< MboII          >< AfaI          >< BsaB1Alw26I ><
GGAAGAAGAA ATTGATGAAA CCTGTGAACA TGAGTACGGT ACAGAGGATG ATTATCAAGG TCTCCCTCTG
    3090          3100          3110          3120          3130          3140          3150

    >< NlaIV>< PvuII>< XmnI
    >< Eco64I >< Psp5I          >< TthHB8I
    >< MnlI >< DdeI          >< TaqI          >< MnlI          >< MboII
    >< BscBI>< NspBII >< MnlI          >< Ksp632I          >< MboII >< MboII
    >< BanI          >< MnlI          >< EarI          >< BsrI
    >< AccBII >< AluI >< Asp700I          >< Eam1104I >< MboII>< BbsI
GAATTTGGTG CCTCAGCTGA AACAGTTCGA GTTGAGGAAG AAGAAGAGGA AGACTGGCTG GATGATACTA
    3160          3170          3180          3190          3200          3210          3220

                                >< Tru9I
    >< FokI          >< MseI          >< Eco57I
    >< DdeI          >< BsrI>< MboII BsrI ><
CTGAGCAATC AGAGATTGAG CCAGAACCAG AACCTACACC TGAAGAACCA GTTAATCAGT TTAAGTGGTTA
    3230          3240          3250          3260          3270          3280          3290

    >< Tru9I          >< MnlI
    >< MseI          >< Tru9I >< HindII>< Tru9I          >< DraIII
    >< DraI          >< MseI          >< HincII>< MseI          >< BspWI
TTTAAAACTT ACTGACAATG TTGCCATTAA ATGTGTTGAC ATCGTTAAGG AGGCACAAAG TGCTAATCCT
    3300          3310          3320          3330          3340          3350          3360

                                >< VneI
                                >< SnaI
                                >< SduI
                                >< NspII
                                >< HgiAI
                                >< Bsp1286I
                                >< BmyI
                                >< ApaLI
    >< HphI          >< NlaIII          >< Alw44I
    >< BbvI          >< Fnu4HI          >< BspMI          >< Alw21I
ATGGTGATTG TAAATGCTGC TAACATACAC CTGAAACATG GTGGTGGTGT AGCAGGTGCA CTCAACAAGG
    3370          3380          3390          3400          3410          3420          3430

                                >< Sau96I
                                >< Pali
                                >< NspIV
                                >< HaeIII
    >< NlaIV          >< Cfr13I

```

FIGURA 13. 9

```

>< Eco64I
>< BscBI
>< BanI
>< AccB1I>< NlaIII
CAACCAATGG TGCCATGCAA AAGGAGAGTG ATGATTACAT TAAGCTAAAT GGCCCTCTTA CAGTAGGAGG
3440 3450 3460 3470 3480 3490 3500

>< BsuRI
>< BsiZI
>< BshI
>< MnlI
>< AluI
>< AsuI
>< MnlI
>< SinI
>< Sau96I
>< NspIV
>< NspHI>< NspHII
>< Eco47I
>< Cfr13I
>< NlaIII
>< BspMI
>< BsiZI
>< Bme18I
>< AvaII MnlI ><
>< DdeI
>< NspI>< AsuI FokI ><
GTCTTGTTTG CTTTCTGGAC ATAATCTTGC TAAGAAAGTGT CTGCATGTTG TTGGACCTAA CCTAAATGCA
3510 3520 3530 3540 3550 3560 3570

>< Tru9I
>< HphI> < MseI
>< Esp4I
>< AluI
>< NdeI
>< AflII>< Fnu4HI
>< BbvI
GGTGAGGACA TCCAGCTTCT TAAGGCAGCA TATGAAAATT TCAATTCACA GGACATCTTA CTTGCACCAT
3580 3590 3600 3610 3620 3630 3640

RsaI ><
Csp6I ><
AfaI ><
>< Eco57I
>< BcgI
TGTGTCAGC AGGCATATTT GGTGCTAAAC CACTTCAGTC TTTACAAGTG TGCGTGCAGA CGGTTCGTAC
3650 3660 3670 3680 3690 3700 3710

>< BsqI
>< BcgI/a
>< BspMI
>< AluI
>< NlaIII
ACAGGTTTAT ATTGCAGTCA ATGACAAAGC TCTTTATGAG CAGGTTGTCA TGGATTATCT TGATAACCTG
3720 3730 3740 3750 3760 3770 3780

>< MnlI
>< RmaI
>< MaeI
>< MnlI
>< Eco57I
>< BscBI
>< TfiI
>< MboII
>< HinfI
>< DdeI
AAGCCTAGAG TGGAAGCACC TAAACAAGAG GAGCCACCAA ACACAGAAGA TTCCAAAACCT GAGGAGAAAT
3790 3800 3810 3820 3830 3840 3850

>< Tru9I
>< StuI
>< Pali
>< MseI
>< MnlI
>< MaeIII
>< HaeIII
>< EcoO65I
>< Eco147I
>< Eco91I
>< BsuRI
>< BstXI ><
>< RsaI
>< Csp6I
>< AfaI
>< TthHB8I
>< TaqI
>< BshI
>< BstPI
>< AatI
>< BstEII
CTGTCGTACA GAAGCCTGTC GATGTGAAGC CAAAAATTAA GGCCTGCATT GATGAGGTTA CCACAACACT
3860 3870 3880 3890 3900 3910 3920

TfiI ><
NlaIII ><
HinfI ><
>< DdeI
>< EcoRV
>< HindIII

```

FIGURA 13. 10

```

>< BsrI      >< MboII      >< MaeIII      >< Eco32I      >< AluI
GGAAGAAACT  AAGTTTCTTA  CCAATAAGTT  ACTCTTGTTT  GCTGATATCA  ATGGTAAGCT  TTACCATGAT
3930        3940        3950        3960        3970        3980        3990

      >< NspI
      >< NspHI
      >< NlaIII
    >< MnlI      >< SfaNI
                  > < EcoNI
      >< DdeI      >< DdeI      >< MboII >< BslI      > < NlaIII
    >< DdeI      >< BfrI      >< HphI      >< BsiYI      >< FokI
TCTCAGAACA  TGCTTAGAGG  TGAAGATATG  TCTTTCCTTG  AGAAGGATGC  ACCTTACATG  GTAGGTGATG
4000        4010        4020        4030        4040        4050        4060

    >< SpeI
    >< RmaI
    >< MaeI      >< EcoRV>< HphI      >< SfaNI
    >< HphI      >< Eco32I      >< MnlI      >< DdeI
TTATCACTAG  TGGTGATATC  ACTTGTTGTTG  TAATACCCCTC  CAAAAGGCT  GGTGGCACTA  CTGAGATGCT
4070        4080        4090        4100        4110        4120        4130

                                >< ScrFI
                                >< RsaI
                                >< MvaI
                                >< EcoRII
                                >< Ecl136I
                                >< DsaV
                                >< Csp6I >< EcoNI
                                >< BstOI
                                >< BstNI
                                >< BsiLI
                                >< BsaJI
                                >< BsaAI >< BslI
                                >< MaeII>< ApyI
    >< AluI      >< BsrI      >< AfaI >< BsiYI
CTCAAGAGCT  TTGAAGAAAG  TGCCAGTTGA  TGAGTATATA  ACCACGTACC  CTGGACAAGG  ATGTGCTGCT
4140        4150        4160        4170        4180        4190        4200

                                >< Tru9I
                                >< MseI
                                >< DdeI >< Esp4I      >< RsaI
    >< MnlI      >< BspWI      >< Csp6I
    >< FokI      >< AluI      >< AflIII      >< Eco57I >< AfaI
TATACACTTG  AGGAAGCTAA  GACTGCTCTT  AAGAAATGCA  AATCTGCATT  TTATGTACTA  CCTTCAGAAG
4210        4220        4230        4240        4250        4260        4270

                                >< ScrFI
                                >< MvaI
                                >< EcoRII
                                >< Ecl136I
                                >< XmnI      >< NlaIII ><
    > < Ksp632I      >< RmaI      >< DsaV      >< Ksp632I ><
    > < EarI > < TfiI>< MboII      >< BstOI      >< EarI
    > < Eam1104I      >< MaeI      >< BstNI      >< Eam1104I ><
    > < DdeI > < HinfI      >< BsiLI      >< BsmAI ><
    >< BspWI      >< Asp700I      >< ApyI      >< Alw26I ><
CACCTAATGC  TAAGGAAGAG  ATTCTAGGAA  CTGTATCCTG  GAATTTGAGA  GAAATGCTTG  CTCATGCTGA
4280        4290        4300        4310        4320        4330        4340

    >< VspI      >< Zsp2I
    >< Tru9I      >< Ppu10I
    >< MseI      >< NsiI
    >< MboII      >< NlaIII      >< FokI
    >< Eco57I      >< Mph1103I      >< FokI

```

FIGURA 13. 11

ES 2 396 127 T3

```

                >< AsnI          >< EcoT22I          >< BspWI
                >< AseI          >< AvaIII          >< BglI          >< MaeII
AGAGACAAGA AAATTAATGC CTATATGCAT GGATGTTAGA GCCATAATGG CAACCATCCA ACGTAAGTAT
    4350          4360          4370          4380          4390          4400          4410

                >< SfaNI
                >< Tru9I          > < HindII          >< TfiI          >< SpeI
                >< MseI          > < HincII>< MboII          >< RmaI
                >< MnlI          >< DrdI >< HinfI          >< MaeI
AAAGGAATTA AAATTCAAGA GGGCATCGTT GACTATGGTG TCCGATTCTT CTTTATACT AGTAAAGAGC
    4420          4430          4440          4450          4460          4470          4480

                >< MaeIII
>< SfcI          >< Fnu4HI          >< MunI
>< AluI          >< AluI          >< AciI          MaeIII ><
CTGTAGCTTC TATTATTACG AAGCTGAACT CTCTAAATGA GCCGCTTGTC ACAATGCCAA TTGGTTATGT
    4490          4500          4510          4520          4530          4540          4550

                >< ThaI
                >< MvnI
                >< MboII
                >< HinPII
>< HinPII
>< Hin6I
>< Hin6I
>< HhaI
>< HhaI
>< Tru9I          >< HhaI
>< NlaIII          >< Fnu4HI
>< MseI          >< CfoI
>< MnlI          >< CfoI
>< Ksp632I          >< BstUI
>< EarI          >< BssHII>< BspWI          >< Tru9I
>< Eam1104I          >< Bsp50I          >< MseI
>< BbvI          >< AccII          >< AluI          HphI ><
GACACATGGT TTTAATCTTG AAGAGGCTGC GCGCTGTATG CGTTCCTTA AAGCTCCTGC CGTAGTGTCA
    4560          4570          4580          4590          4600          4610          4620

                >< MaeIII
>< SfaNI          >< AlwNI          >< MnlI >< MnlI>< DdeI
GTATCATCAC CAGATGCTGT TACTACATAT AATGGATACC TCACTTCGTC ATCAAAGACA TCTGAGGAGC
    4630          4640          4650          4660          4670          4680          4690

                >< SinI
                >< Sau96I
                >< NspIV
                >< NspHII
>< SduI          >< Eco47I
>< NspII          >< Cfr13I
>< HgiAI          >< BsiZI
>< Bsp1286I          >< Bme18I          >< RsaI
>< BmyI          >< AvaII          >< Csp6I
>< Alw21I          >< AsuI          >< AfaI
ACTTTGTAGA AACAGTTTCT TTGGCTGGCT CTTACAGAGA TTGGTCCCTAT TCAGGACAGC GTACAGAGTT
    4700          4710          4720          4730          4740          4750          4760

                > < TthHB8I
                > < TaqI
                >< SduI
                >< Van91I >< NspII
                >< Tru9I          >< RsaI >< PflMI >< Eco24I
                >< MseI          >< HphI >< BslI >< Bsp1286I
                >< Esp4I          >< Csp6I >< BsiYI >< BmyI >< GsuI ><

```

FIGURA 13. 12

```

                >< AflIII >< MaeIII                >< AfaI >< AccB7I >< BanIIBpmI ><
AGGTGTTGAA TTTCTTAAGC GTGGTGACAA AATTGTGTAC CACACTCTGG AGAGCCCCGT CGAGTTTCAT
    4770          4780          4790          4800          4810          4820          4830

                                >< Tru9I
                                >< P1eI >< EcoNI
                                >< MnlI >< BslI
                                >< BsmAI >< BsiYI
    >< MnlI                >< HphI                >< HinfI>< Alw26I>< AciI >< MseI
CTTGACGGTG AGGTTCTTTC ACTTGACAAA CTAAAGAGTC TCTTATCCCT GCGGGAGGTT AAGACTATAA
    4840          4850          4860          4870          4880          4890          4900

                                >< AluI                >< NdeI
AAGTGTTTAC AACTGTGGAC AACACTAATC TCCACACACA GCTTGTGGAT ATGTCTATGA CATATGGACA
    4910          4920          4930          4940          4950          4960          4970

    >< SinI
    >< Sau96I
    >< NspIV
    >< NspHII
    >< Eco47I
    >< Cfr13I
    >< BsiZI
    >< Bme18I
    >< AvaII
    >< AsuI
                                NlaIII ><
                                >< NlaIII
                                > < MnlI
    >< MaeIII >< Tru9I >< MnlI
    >< FokI >< MseI >< BspHI
GCAGTTTGGT CCAACATACT TGGATGGTGC TGATGTTACA AAAATTAAAC CTCATGTAAA TCATGAGGGT
    4980          4990          5000          5010          5020          5030          5040

                                > < TthHB8I
                                > < TaqI
    >< RsaI
    > < RmaI                >< SnaBI                >< ScaI
    > < MaeI                >< MaeII >< HindIII >< RsaI
    >< Csp6I                >< Eco105I                >< Csp6I
    >< AfaI                >< BsaAI >< AluI >< AfaI
AAGACTTTCT TTGTACTACC TAGTGATGAC ACACTACGTA GTGAAGCTTT CGAGTACTAC CATACTCTTG
    5050          5060          5070          5080          5090          5100          5110

    >< RsaI
    >< NspI
    >< NspHI
    >< NlaIII
    > < Csp6I                >< Tru9I                MnlI >
    >< AflIII >< MseI                BslI ><
    >< AfaI >< DraI                BsiYI ><
ATGAGAGTTT TCTTGGTAGG TACATGTCTG CTTTAAACCA CACAAAGAAA TGGAAATTC CTCAAGTTGG
    5120          5130          5140          5150          5160          5170          5180

    >< Tru9I >< Tru9I                >< RmaI
    >< MseI >< MseI                >< MunI >< MaeI                AluI >
TGGTTTAACT TCAATTAAAT GGGCTGATAA CAATTGTTAT TTGTCTAGTG TTTTATTAGC ACTTCAACAG
    5190          5200          5210          5220          5230          5240          5250

                                >< SfaNI
                                >< SduI
                                >< NspII
                                >< Eco24I
                                >< Bsp1286I
                                >< BmyI                HphI >
                                >< BbvI Fnu4HI ><
                                >< BanII >< BspWI
    >< MnlI

```

FIGURA 13. 13

```

CTTGAAGTCA AATTCAATGC ACCAGCACTT CAAGAGGCTT ATTATAGAGC CCGTGCTGGT GATGCTGCTA
5260 5270 5280 5290 5300 5310 5320

>< VneI
>< SnoI
    >< SduI
    >< NspII
    >< HgiAI
    >< Bsp1286I
    >< BmyI
>< ApaLI
>< Alw44I
    >< Alw21I
    >< AluI
    MboII ><
    >< HphI
ACTTTTGTGC ACTCATACTC GCTTACAGTA ATAAAACTGT TGGCGAGCTT GGTGATGTCA GAGAACTAT
5330 5340 5350 5360 5370 5380 5390

    > < SphI
    > < PaeI
    > < NspI
    > < NspHI >< TfiI
    >< Tru9I
    >< SfcI > < NlaIII>< HinfI
    >< MseI
GACCCATCTT CTACAGCATG CTAATTTGGA ATCTGCAAAG CGAGTTCTTA ATGTGGTGTG TAAACATTGT
5400 5410 5420 5430 5440 5450 5460

    >< RsaI
    >< Tru9I
    > < Csp6I
    Esp4I >
    >< MseI
    >< AluI
    >< AfaI
    AflIII >
GGTCAGAAAA CTACTACCTT AACGGGTGTA GAAGCTGTGA TGTATATGGG TACTCTATCT TATGATAATC
5470 5480 5490 5500 5510 5520 5530

    >< RsaI
    >< MboII
    >< RmaIHinfI ><
    >< Csp6I
>< Tru9I
    >< SfaNI
    >< MaeI >< BbsI
>< MseI
    >< NlaIII
    >< AfaI
TTAAGACAGG TGTTTCCATT CCATGTGTGT GTGGTCTGTA TGCTACACAA TATCTAGTAC AACAGAGTC
5540 5550 5560 5570 5580 5590 5600

    >< RsaI
    >< PleI
    > < DdeI
    >< Csp6I
    >< BsgI
    >< BspWI >< BspMI
    >< AfaI
TTCTTTTGGT ATGATGTCTG CACCACCTGC TGAGTATAAA TTACAGCAAG GTACATTCTT ATGTGCGAAT
5610 5620 5630 5640 5650 5660 5670

    >< RsaI
    >< Eco3I
    >< DdeI
    > < MaeIII
    >< BsmAI
    >< Csp6I
    >< BsaI
    MnlI ><
    >< AfaI >< BsrI
    >< Alw26I
    HphI >
GAGTACACTG GTAACATCA GTGTGGTCAT TACACTCATA TAACTGCTAA GGAGACCCTC TATCGTATTG
5680 5690 5700 5710 5720 5730 5740

    >< SstI
    >< SinI
    >< SduI
    >< Sau96I
    >< SacI
    >< NspIV
    >< NspHII
    >< HgiAI
    > < RsaI
    >< MaeIII
    >< Eco24I
    >< Eco47I
    >< Ecl136II
    >< Cfr13I
    >< Bsp1286I
    >< Bsi2I
    >< BmyI
    >< Bmel8I

```

FIGURA 13. 14

ES 2 396 127 T3

```

    >< BanII
    >< Alw21I
    >< AluI
ACGGAGCTCA CCTTACAAAG ATGTCAGAGT ACAAAGGACC AGTGACTGAT GTTTTCTACA AGGAAACATC
    5750          5760          5770          5780          5790          5800          5810

    >< TthHB8I
    >< TaqI >< MaeIII
TTCACTACA ACCATCAAGC CTGTGTCGTA TAAACTCGAT GGAGTTACTT ACACAGAGAT TGAACCAAAA
    5820          5830          5840          5850          5860          5870          5880

    >< RsaI
    >< Csp6I
    >< SfcI >< BbvI
    >< FokI
    >< Fnu4HI
    >< AfaI
TTGGATGGGT ATTATAAAAA GGATAATGCT TACTATACAG AGCAGCCTAT AGACCTTGTA CCAACTCAAC
    5890          5900          5910          5920          5930          5940          5950

    Tru9I ><
    SmaI ><
    MseI ><
    > < NspI
    > < NspHI
    > < NlaIII
    >< AflIII
    >< BsaBI ><
CATTACCAAA TGCGAGTTTT GATAATTTCA AACTCACATG TTCTAACACA AAATTTGCTG ATGATTTAAA
    5960          5970          5980          5990          6000          6010          6020

    >< MboII
    >< AluI
    >< AluI>< MaeIII
TCAAATGACA GGCTTCACAA AGCCAGCTTC ACGAGAGCTA TCTGTCACAT TCTTCCCAGA CTTGAATGGC
    6030          6040          6050          6060          6070          6080          6090

    >< SfcI
GATGTAGTGG CTATTGACTA TAGACACTAT TCAGCGAGTT TCAAGAAAGG TGCTAAATTA CTGCATAAGC
    6100          6110          6120          6130          6140          6150          6160

    >< Tru9I
    >< ScrFI
    >< MvaI
    >< MseI
    >< EcoRII
    >< Ecl136I
    >< DsaV
    >< BstOI
    >< BstNI
    >< BsiLI
    >< MunI
    >< BstXI
    >< ApyI
    >< MaeII
    >< BstXI
    >< DraIII
    >< MaeII ><
    >< BbsI
CAATTGTTTG GCACATTAAC CAGGCTACAA CCAAGACAAC GTTCAAACCA AACACTTGGT GTTTACGTTG
    6170          6180          6190          6200          6210          6220          6230

    > < RsaI
    >< Csp6I
    > < AfaI>< BsrI
    >< MboII ><
    >< BbsI
TCTTTGGAGT ACAAAGCCAG TAGATACTTC AAATTCATTT GAAGTTCTGG CAGTAGAAGA CACACAAGGA
    6240          6250          6260          6270          6280          6290          6300

    >< HindII
    >< HincII
    >< MnlI
    >< MboII
    >< Eco57I
ATGGACAATC TTGCTTGTGA AAGTCAACAA CCCACCTCTG AAGAAGTAGT GGAAAATCCT ACCATACAGA
    6310          6320          6330          6340          6350          6360          6370

```

FIGURA 13. 15

ES 2 396 127 T3

```

                >< MaeIII                                >< Tru9I
                >< MaeII                                >< MseI
AGGAAGTCAT AGAGTGTGAC GTGAAAAC TA CCGAAGTTGT AGGCAATGTC ATACTTAAAC CATCAGATGA
    6380         6390         6400         6410         6420         6430         6440

                .
                >< XhoII
                >< Sau3AI
                >< NlaIII
                >< NdeII
                >< MflI
                >< MboI
                >< DpnII
                >< DpnI
                >< BstYI
                >< BspAI
    >< Tru9I
    >< MseI
                >< BspHI >< BspI43I>< Fnu4HI
                > < MaeIII >< MnlI >< BbvI >< AlwI
AGGTGTTAAA GTAACACAAG AGTTAGGTCA TGAGGATCTT ATGGCTGCTT ATGTGGAAAA CACAAGCATT
    6450         6460         6470         6480         6490         6500         6510

                >< SauI
                >< RmaI
                >< MstII
                >< MaeI
                >< Eco81I
                >< DdeI
                >< CvnI
                >< Bsu36I
                >< Bse21I
                >< BfrI> < Tru9I
    >< Tru9I
    >< MseI
                >< AluI
                >< AxyI> < MseI>< MunI
                >< AocI >< DraI >< BbvI Fnu4HI >< NlaIII
ACCATTAAGA AACCTAATGA GCTTTCACTA GCCTTAGGTT TAAAAACAAT TGCCACTCAT GGTATTGCTG
    6520         6530         6540         6550         6560         6570         6580

                >< VspI >< StyI
                >< Tru9I >< EcoT14I
                >< MseI >< Eco130I
                >< AsnI >< BssT1I
                >< AseI >< BsaJI
                >< DdeI
                >< BslI
                >< BsiYI
                >< BfrI >< Fnu4HI
CAATTAATAG TGTTCCTTGG AGTAAATTT TGGCTTATGT CAAACCATTG TTAGGACAAG CAGCAATTAC
    6590         6600         6610         6620         6630         6640         6650

                >< HinP1I
                >< Hin6I
                >< HhaI
                >< DdeI
                >< BbvI >< CfoI
                >< Tru9I
                >< MaeII>< MseI
                >< DraIII
                >< AflIII
AACATCAAAT TCGGCTAAGA GATTAGCACA ACGTGTGTTT AACAATTATA TGCCTTATGT GTTTACATTA
    6660         6670         6680         6690         6700         6710         6720

                >< RsaI
                >< Csp6I
                >< MunI >< AfaI
                >< RsaI>< XbaI
                >< Csp6I >< RmaI
                >< AfaI >< MaeI
                >< AluI
TTGTTCCAAT TGTGTACTTT TACTAAAAGT ACCAATTCTA GAATTAGAGC TTCCTACCT ACAACTATTG
    6730         6740         6750         6760         6770         6780         6790

                >< VspI
                >< Tru9I
                >< NaeI
                >< MspI
                >< MseI

```

FIGURA 13. 16

```

                                >< HpaII
                                >< HapII
                                >< Cfr10I >< FokI
                                >< AsnI
                                >< AseI>< HphI>< MaeIII
                                >< Tru9I
                                >< MseI
                                >< SfaNI
                                >< DdeI
                                >< BfrI
                                >< BbvI
                                >< SduI
                                >< NspII
                                >< HgiAI
                                >< BspI286I
                                >< BmyI
                                >< Alw21I
                                >< RsaI
                                >< Csp6I
                                >< Fnu4HI
                                >< AfaI
                                >< RsaI ><
                                >< MseI ><
                                >< Fnu4HI
                                >< MaeIII
                                >< MaeII
                                >< BbvI >
                                >< TfiI
                                >< MamI
                                >< HinfI
                                >< BsiBI
                                >< XmnI>< MaeIII
                                >< AluI >
                                >< PleI>< HinfI
                                >< BsaBI >< AluI
                                >< Asp700I
                                >< AfaI ><
                                >< Pali
                                >< Nsp8II
                                >< HaeIII
                                >< GdiII
                                >< Fnu4HI
                                >< EaeI
                                >< DdeI
                                >< BsuRI
                                >< RmaI
                                >< BshI >< BslI
                                >< MaeI
                                >< AciI>< BsiYI
                                >< BspMI
                                >< AluI
                                >< RmaI
                                >< MaeI
                                >< NlaIV
                                >< Eco64I
                                >< RsaI >< BscBI
                                >< Csp6I >< BanI
                                >< AfaI>< AccBII
                                >< NlaIII
                                >< RsaI ><
                                >< MboII
                                >< MamI ><
                                >< Csp6I ><
                                >< BsiBI ><
                                >< BsaBI ><
                                >< AfaI ><

```

CTAAAAATAG TGTTAAGAGT GTTGCTAAAT TATGTTTGGA TGCCGGCATT AATTATGTGA AGTCACCCAA
6800 6810 6820 6830 6840 6850 6860

ATTTTCTAAA TTGTTACAAA TCGCTATGTG GCTATTGTTG TTAAGTATTT GCTTAGGTTC TCTAATCTGT
6870 6880 6890 6900 6910 6920 6930

GTAAGTCTG CTTTTGGTGT ACTCTTATCT AATTTTGGTG CTCCTTCTTA TTGTAATGGC GTTAGAGAAT
6940 6950 6960 6970 6980 6990 7000

TGTATCTTAA TTCGTCTAAC GTTACTACTA TGGATTCTCG TGAAGGTTCT TTTCCTTGCA GCATTTGTTT
7010 7020 7030 7040 7050 7060 7070

AAGTGGATTA GACTCCCTTG ATTCTTATCC AGCTCTTGAA ACCATTTCAGG TGACGATTTC ATCGTACAAG
7080 7090 7100 7110 7120 7130 7140

CTAGACTTGA CAATTTTAGG TCTGGCCGCT GAGTGGGTTT TGGCATATAT GTTGTTCACA AAATTCCTTT
7150 7160 7170 7180 7190 7200 7210

ATTATTAGG TCTTTCAGCT ATAATGCAGG TGTTCTTTGG CTATTTTGCT AGTCATTTC TCAGCAATTC
7220 7230 7240 7250 7260 7270 7280

FIGURA 13. 17

```

TTGGCTCATG TGGTTTATCA TTAGTATTGT ACAAATGGCA CCCGTTCTG CAATGGTTAG GATGTACATC
    7290          7300          7310          7320          7330          7340          7350

                                TthHB8I ><
                                >< TaqI
                                MnlI ><
                                >< NdeI
                                >< Ksp632I
                                >< EarI
                                >< Eam1104I>< AluI>< MboII >< NlaIII Eam1104I ><
                                >< FokI
                                >< MboII EarI ><
                                >< FokI
                                >< Eam1104I>< AluI>< MboII >< NlaIII Eam1104I ><
TTCTTTGCTT CTTTCTACTA CATATGGAAG AGCTATGTTT ATATCATGGA TGGTTGCACC TCTTCGACTT
    7360          7370          7380          7390          7400          7410          7420

                                XhoII ><
                                Sau3AI ><
                                NlaIII ><
                                NdeII ><
                                MflI ><
                                MboI ><
                                >< Thal
                                >< MvnI
                                >< Ksp632I
                                >< EarI
                                >< Eam1104I
                                >< HinPII
                                >< MluI
                                >< BstUI
                                >< Bsp50I >< RsaI
                                >< DpnII ><
                                >< HhaI
                                >< Bsp6I
                                >< BstYI ><
                                >< NlaIII
                                >< CfoI
                                >< AflIII >< Csp6I
                                >< Tru9I BspAI ><
                                >< BspWI >< BspWI
                                >< AccII >< AfaI
                                >< MseI BglII ><
GCATGATGTG CTATAAGCGC AATCGTGCCA CACGCGTTGA GTGTACAACT ATTGTTAATG GCATGAAGAG
    7430          7440          7450          7460          7470          7480          7490

                                >< Pali
                                >< HaeIII
                                >< DsaI
                                >< MunI
                                >< MboII
                                >< BsuRI
                                >< MaeIII ><
                                >< DpnI
                                >< BshI
                                >< MunI
                                >< BsmAI ><
                                >< Bsp143I
                                >< MnlI
                                >< BsaJI >< PleI>< HinfI
                                >< Alw26I ><
ATCTTTCTAT GTCTATGCAA ATGGAGGCCG TGGCTTCTGC AAGACTCACA ATTGGAATTG TCTCAATTGT
    7500          7510          7520          7530          7540          7550          7560

                                >< RsaI
                                >< Tru9I ><
                                >< Csp6I
                                >< MseI ><
                                >< BsrI
                                >< GsuI >< MaeIIIDraI ><
                                >< AfaI
                                >< BpmI
                                >< BsrI
GACACATTTT GCACTGGTAG TACATTCATT AGTGATGAAG TTGCTCGTGA TTTGTCACTC CAGTTTAAAA
    7570          7580          7590          7600          7610          7620          7630

                                >< Thal
                                >< MvnI
                                >< HphI
                                >< HinPII ><
                                >< HinPII
                                >< Hin6I
                                >< Hin6I
                                >< HhaI ><
                                >< HhaI
                                >< CfoI ><
                                >< CfoI
                                >< BstUI
                                >< BssHII
                                >< Bsp50I ><
                                >< BsrI
                                >< AccII
GACCAATCAA CCCTACTGAC CAGTCATCGT ATATTGTTGA TAGTGTTGCT GTGAAAAATG GCGCGCTTCA
    7640          7650          7660          7670          7680          7690          7700

```

FIGURA 13. 18

```

                                >< FokI
                                >< BsmAI
                                >< Alw26I
                                >< MnlI
                                >< AciI
CCTCTACTTT GACAAGGCTG GTCAAAAGAC CTATGAGAGA CATCCGCTCT CCCATTTTGT CAATTTAGAC
7710      7720      7730      7740      7750      7760      7770

                                >< VspI
                                >< Tru9I
                                >< MseI
                                >< AsnI
                                >< AseI
                                >< BcgI/a
> < AluI
AATTTGAGAG CTAACAACAC TAAAGGTTCA CTGCCTATTA ATGTCATAGT TTTTGATGGC AAGTCCAAAT
7780      7790      7800      7810      7820      7830      7840

                                >< SfcI
                                >< PvuII
                                >< RsaI
                                >< Psp5I
                                >< Csp6I
                                >< NspBII
                                >< PleI
                                >< HinfI
                                >< DdeI
                                >< BcgI
                                >< AfaI
                                >< AluI
GCGACGAGTC TGCTTCTAAG TCTGCTTCTG TGTACTACAG TCAGCTGATG TGCCAACCTA TTCTGTTGCT
7850      7860      7870      7880      7890      7900      7910

                                TthHB8I ><
                                TaqI ><
                                SalI ><
                                RtrI ><
                                HindII >
                                >< ScaI
                                >< RsaI
                                >< Tru9I
                                >< Csp6I
                                >< SfaNI
                                >< Eco57I
                                >< AluI
                                >< MaeII
                                >< AfaI
                                >< MseI
                                >< AccI
TGACCAAGCT CTTGTATCAG ACGTTGGAGA TAGTACTGAA GTTTCCGTTA AGATGTTTGA TGCTTATGTC
7920      7930      7940      7950      7960      7970      7980

                                >< Tru9I
                                >< MseI
                                >< Esp4I
                                >< SfcI
                                >< AflIII
                                >< BspWI
                                >< AluI
GACACCTTTT CAGCAACTTT TAGTGTTCCT ATGGAAAAAC TTAAGGCACT TGTGCTTACA GCTCACAGCG
7990      8000      8010      8020      8030      8040      8050

                                >< PvuII
                                >< Psp5I
                                >< NspBII
                                >< Fnu4HI
                                >< AluI
                                >< BbvI
AGTTAGCAAA GGGTGTAGCT TTAGATGGTG TCCTTTCTAC ATTCGTGTCA GCTGCCCCGAC AAGGTGTTGT
8060      8070      8080      8090      8100      8110      8120

                                MaeIII ><
                                >< HindII
                                >< BsmAI
                                >< DdeI
                                >< HincII
                                >< FokI
                                >< Alw26I
                                >< BfrI
TGATACCGAT GTTGACACAA AGGATGTTAT TGAATGTCTC AAACTTTCAC ATCACTCTGA CTTAGAAGTG
8130      8140      8150      8160      8170      8180      8190

                                >< XhoII
                                Sau3AI ><
                                >< NdeII
                                >< MflI
                                >< MboI
                                >< NlaIII
                                >< HgaI
                                >< HinfI
                                >< DpnII
                                DpnI ><

```

FIGURA 13. 19

```

                                Bsp143I ><
                                >< BsaHI >< BstYI
                                >< BbiII >< BspAI
                                >< AclI >< BglII
                                >< MaeIII>< HphI
                                >< MaeIII >< HphI >< NlaIII
ACAGGTGACA GTTGTAAACA TTTTCATGCTC ACCTATAATA AGGTTGAAAA CATGACGCCC AGAGATCTTG
      8200      8210      8220      8230      8240      8250      8260

                                >< NspI
                                >< NspHI
                                >< NlaIII
>< HinPII
>< Hin6I
>< HhaI
>< CfoI
                                >< BspWI >< MaeIII
GCGCATGTAT TGA CTGTAAT GCAAGGCATA TCAATGCCCA AGTAGCAAAA AGTCACAATG TTTCACATCAT
      8270      8280      8290      8300      8310      8320      8330

                                >< NspI
                                >< NspHI
                                >< NlaIII
                                >< PvuII
                                >< Psp5I
                                >< Eam1105I
                                >< NspBII
                                >< BbvI
                                >< Fnu4HI
                                >< AflIII
                                >< AluI >< BbvI >< Fnu4HI
CTGGAATGTA AAAGACTACA TGTCTTTATC TGAACAGCTG CGTAAACAAA TTCGTAGTGC TGCCAAGAAG
      8340      8350      8360      8370      8380      8390      8400

                                >< RmaI
                                >< MaeI >< Eam1105I
>< MboII
AACAAATAC CTTT TAGACT AACTTGTGCT ACAACTAGAC AGGTTGTCAA TGTCAATACT ACTAAAATCT
      8410      8420      8430      8440      8450      8460      8470

                                >< Tru9I
                                >< Pali
                                >< MseI
                                >< HaeIII
                                >< ScaI
                                >< Esp4I
                                >< RsaI >< Tru9I
                                >< BsuRI
                                >< Csp6I >< MseI
                                >< BshI
                                >< AfaI >< DraI
                                >< AflIII >< BbvI
CACTCAAGGG TGCTAAGATT GTTAGTACTT GTTTTAACT TATGCTTAAG GCCACATTAT TGTGCGTTCT
      8480      8490      8500      8510      8520      8530      8540

                                >< RsaI
                                >< Csp6I
                                >< BsrI
                                >< NlaIII
                                >< Fnu4HI
                                >< AfaI
                                >< MaeIII
TGCTGCATTG GTTTGTTATA TCGTTATGCC AGTACATACA TTGTCAATCC ATGATGGTTA CACAAATGAA
      8550      8560      8570      8580      8590      8600      8610

                                >< MaeIII
                                >< MaeIII
                                >< FokI
ATCATTGGTT ACAAAGCCAT TCAGGATGGT GTCACGCTG ACATCATTTT TACTGATGAT TGTTTGTCAA
      8620      8630      8640      8650      8660      8670      8680

                                SfcI >
                                Fnu4HI ><
                                BbvI ><
                                >< NspI
                                >< NspHI
                                >< NlaIII
                                >< RgaI
                                >< BstXI
                                >< BbvI
                                >< AluI
ATAAACATGC TGTTTGTGAC GCATGGTTTA GCCAGCGTGG TGTTTCATAC AAAAATGACA AAAGTGCCCC
      8690      8700      8710      8720      8730      8740      8750

```

FIGURA 13. 20

```

                                >< ScrFI
                                >< ScrFI >< RsaI
                                >< MvaI >< MspI
                                >< EcoRII >< HpaII
                                >< Ecl136I>< NciI
                                >< DsaV >< HapII
                                >< BstOI>< DsaV
                                >< BstNI >< Csp6I
                                >< BsiLI >< BcnIDdeI ><
                                >< ApyI >< AfaI
                                >< Fnu4HI
                                >< AluI
TGTAGTAGCT GCTATCATT CAAGAGAGAT TGGTTTCATA GTGCCTGGCT TACCGGGTAC TGTGCTGAGA
8760 8770 8780 8790 8800 8810 8820

                                >< MaeIII >< HphI >< MnlI >< BspWI
GCAATCAATG GTGACTTCTT GCATTTTCTA CCTCGTGTTC TTAGTGCTGT TGGCAACATT TGCTACACAC
8830 8840 8850 8860 8870 8880 8890

                                Tru9I >
                                SfaNI ><
                                >< RsaI
                                MseI >
                                >< BspWI >< Fnu4HI >< Csp6I
                                >< BbvI>< MnlI >< DdeI >< AfaI
CTTCCAACT CATTGAGTAT AGTGATTTTG CTACCTCTGC TTGCGTCTT GCTGCTGAGT GTACAATTTT
8900 8910 8920 8930 8940 8950 8960

                                >< RmaI
                                >< MnlI
                                >< FokI
                                >< MaeI
TAAGGATGCT ATGGGCAAAC CTGTGCCATA TTGTTATGAC ACTAATTTGC TAGAGGGTTC TATTTCTTAT
8970 8980 8990 9000 9010 9020 9030

                                ScrFI >
                                MvaI >
                                MnlI ><
                                EcoRII ><
                                Ecl136I >
                                DsaV ><
                                BstOI >
                                >< NlaIV BstNI >
                                >< FokI BsiLI >
                                >< BscBI ApyI >
                                >< AluI
AGTGAGCTTC GTCCAGACAC TCGTTATGTG CTATGGATG GTTCCATCAT ACAGTTTCTT AACACTTACC
9040 9050 9060 9070 9080 9090 9100

                                >< RsaI
                                >< SfcI >< NspI
                                >< ScaI >< NspHI
                                >< RsaI >< NlaIII
                                >< MaeIII >< Csp6I >< NlaIII
                                >< GsuI >< AfaI >< Csp6I
                                >< BpmI >< DdeI >< AccI >< AfaI
TGGAGGGTTC TGTTAGAGTA GTAACAACCTT TTGATGCTGA GTACTGTAGA CATGGTACAT GCGAAAGGTC
9110 9120 9130 9140 9150 9160 9170

                                >< SstI
                                >< SduI
                                >< SacI
                                NspII ><
                                HgiAI ><
                                Eco24I ><
                                Bsp1286I ><

```

FIGURA 13. 21

```

                                                    Ecl136II ><>< BmyI
                                                    BanII ><
                                                    Alw21I ><
                                                    >< Tru9I
                                                    >< MseI
                                                    >< AluI
AGAAAGTAGGT ATTTGCCTAT CTACCAGTGG TAGATGGGTT CTTAATAATG AGCATTACAG AGCTCTATCA
    9180          9190          9200          9210          9220          9230          9240

                                                    >< TfiI
                                                    >< HinfI >< AluI
                                                    >< MnlI
>< SfaNI
GGAGTTTCT GTGGTGTGA TCGGATGAAT CTCATAGCTA ACATCTTTAC TCCTCTGTG CAACCTGTGG
    9250          9260          9270          9280          9290          9300          9310

                                                    >< MaeIII
                                                    HphI ><
>< Eco57I
GTGCTTTAGA TGTGTCTGCT TCAGTAGTGG CTGGTGGTAT TATTGCCATA TTGGTGACTT GTGCTGCCTA
    9320          9330          9340          9350          9360          9370          9380

                                                    >< RsaI
                                                    >< Csp6I >< NlaIII
                                                    >< MaeII
                                                    >< BbvI >< Fnu4HI
                                                    >< AflIII
                                                    >< AfaI>< HphI >< BspWI
CTACTTTATG AAATTCAGAC GTGTTTTTGG TGAGTACAAC CATGTTGTTG CTGCTAATGC ACTTTTGTTC
    9390          9400          9410          9420          9430          9440          9450

                                                    >< RsaI
                                                    >< NlaIV
                                                    >< KpnI
>< Eco64I
>< Csp6I
>< BscBI
>< Asp718
>< BanI >< AluI
>< AfaI
>< AccBI
>< Acc65I
>< RsaI
>< Csp6I
>< AfaI >< HphI
ACTTGTACTT GACATTCTAT TTCACCAATG ATGTTTCATT CTTGGCTCAC CTTCAATGGT TTGCCATGTT
    9530          9540          9550          9560          9570          9580          9590

                                                    >< HphI
                                                    NlaIII ><
TTCTCCTATT GTGCCTTTT GGATAACAGC AATCTATGTA TTCTGTATTT CTCTGAAGCA CTGCCATTGG
    9600          9610          9620          9630          9640          9650          9660

                                                    >< TthHB8I
                                                    >< RsaI
                                                    >< MnlI
                                                    >< MnlI
                                                    >< Csp6I
>< Tru9I
>< MseI >< DdeI
>< Eco57I >< BfrI >< HinfI >< MseI >< MaeIII >< AfaI Fnu4HI ><
TTCTTTAACA ACTATCTTAG GAAAAGAGTC ATGTTTAATG GAGTTACATT TAGTACCTTC GAGGAGGCTG
    9670          9680          9690          9700          9710          9720          9730

                                                    >< RsaI
                                                    >< Csp6I
                                                    >< BcgI
                                                    >< RsaI
                                                    >< Csp6I
                                                    >< BsmAI

```

FIGURA 13. 22

ES 2 396 127 T3

```

    >< AfaI
CTTTGTGTAC CTTTTTGCTC AACAAGGAAA TGTACCTAAA ATTGCGTAGC GAGACACTGT TGCCACTTAC
    9740          9750          9760          9770          9780          9790          9800

                                >< AfaI          >< Alw26I
                                >< NlaIV
                                >< RsaI          >< DdeI
                                >< Csp6I          >< BscBI
                                >< AfaI          >< BfrI    AluI ><
ACAGTATAAC AGGTATCTTG CTCTATATAA CAAGTACAAG TATTTCAGTG GAGCCTTAGA TACTACCAGC
    9810          9820          9830          9840          9850          9860          9870

                                >< Fnu4HI
                                >< DdeI
                                >< Fnu4HI          >< BfrI
    >< BbvI    >< AluI    >< BbvI          >< DdeI >< AlwNI
TATCGTGAAG CAGCTTGCTG CCACTTAGCA AAGGCTCTAA ATGACTTTAG CAACTCAGGT GCTGATGTTC
    9880          9890          9900          9910          9920          9930          9940

                                >< SfcI          >< BsmI
                                >< EstI          >< BscCI
TCTACCAACC ACCACAGACA TCAATCACTT CTGCTGTTCT GCAGAGTGGT TTTAGGAAAA TGGCATTCCC
    9950          9960          9970          9980          9990          10000          10010

                                >< RsaI
                                >< NlaIII
                                >< MaeIII
                                >< Csp6I          >< Tru9I
                                >< AfaI          >< MseI
GTCAGGCAAA GTTGAAGGGT GCATGGTACA AGTAACCTGT GGAAC TACAA CTCTTAATGG ATTGTGGTTG
    10020          10030          10040          10050          10060          10070          10080

                                XhoII ><
                                Sau3AI ><
                                >< Tru9I    NdeII ><
                                >< NspI      MflI ><
                                >< NspHI      MboI ><
                                >< NlaIII      DpnII ><
                                >< MseI      BstYI ><
                                >< MboII      BspAI ><
                                >< BbsI      BglII ><
GATGACACAG TATACTGTCC AAGACATGTC ATTTGCACAG CAGAAGACAT GCTTAATCCT AACTATGAAG
    10090          10100          10110          10120          10130          10140          10150

                                Pali >
                                MscI >
                                HaeIII >
                                EaeI ><
                                BsuRI >
                                BshI >
                                BalI >
    >< DpnI >< MboII
    >< Bsp143I          >< AluI
ATCTGCTCAT TCGCAAATCC AACCATAGCT TTCTTGTTCA GGCTGGCAAT GTTCAACTTC GTGTTATTGG
    10160          10170          10180          10190          10200          10210          10220

                                >< DdeI> < Tru9I
                                >< BfrI> < MseI          >< DdeI
CCATTCTATG CAAAATTGTC TGCTTAGGCT TAAAGTTGAT ACTTCTAACC CTAAGACACC CAAGTATAAA
    10230          10240          10250          10260          10270          10280          10290

                                >< ScrFI
                                >< MvaI
                                >< EcoRII
                                >< Ecl136I          >< SphI

```

FIGURA 13. 23


```

>< DsaV
>< BstOI
>< BstNI
>< BsiLI
>< ApyI
TTGTGTCGTA TCCAACCTGG TCAAACATTT TCAGTTCTAG CATGCTACAA TGGTTCACCA TCTGGTGTGT
10300 10310 10320 10330 10340 10350 10360

>< PaeI
>< NspI
>< NspHI
>< RmaI >< NlaIII
>< MaeI >< HphI

>< Sau3AI
>< NdeII
>< MboI>< NlaIII
>< DpnII

>< Eco31I
>< BsmAI
>< BsaI>< NlaIII
>< Alw26I
ATCAGTGTGC CATGAGACCT AATCATACCA TTAAAGGTTT TTCCTTAAT GGATCATGTG GTAGTGTGTG
10370 10380 10390 10400 10410 10420 10430

>< Tru9I
>< MseI
>< MseI >< Bsp143I
>< BspAI>< AlwI
>< Zsp2I
>< Ppu10I
>< NsiI>< SfaNI
>< NdeI
>< Mph1103I
>< EcoT22I
>< AvaIII >< AluI >< AfaI ><
>< RsaI >< Csp6I ><
TTTTAACATT GATTATGATT GCGTGTCTTT CTGCTATATG CATCATATGG AGCTTCCAAC AGGAGTACAC
10440 10450 10460 10470 10480 10490 10500

>< SinI
>< Sau96I
>< NspIV
>< NspHII
>< Eco47I
>< Cfr13I
>< BsiZI
>< Bme18I >< HindII
>< AvaII >< HincII
>< AsuI>< BsgI >< BbvI >< BspMI >< AfaI ><
>< RsaI
>< Csp6I>< DdeI
>< AfaI>< BfrI
GCTGGTACTG ACTTAGAAGG TAAATTCTAT GGTCCATTTG TTGACAGACA AACTGCACAG GCTGCAGGTA
10510 10520 10530 10540 10550 10560 10570

>< Tru9I
>< MseI
>< BbvI
>< Fnu4HI
>< HphI ><
CAGACACAAC CATAACATTA AATGTTTTGG CATGGCTGTA TGCTGCTGTT ATCAATGGTG ATAGGTGGTT
10580 10590 10600 10610 10620 10630 10640

>< Tru9I
>< TfiI
>< MseI
>< HphI
>< HinfI
>< Tru9I
>< MseI
>< RsaI
>< Csp6I
>< AfaI
TCTTAATAGA TTCACCACTA CTTTGAATGA CTTTAACCTT GTGGCAATGA AGTACAACCTA TGAACCTTTG
10650 10660 10670 10680 10690 10700 10710

>< SinI
>< Sau96I
>< PssI
>< Psp5II
>< PpuMI
>< NspIV
>< NspHII
>< NlaIV

```

FIGURA 13. 24

```

                                >< EcoO109I
                                >< Eco47I
    >< Sau3AI                    >< DraII
    >< NdeII                     >< Cfr13I
    >< MboI                      >< Bsi2I
    >< DpnII>< NlaIII            >< BscBI
    >< DpnI >< HindII            >< Bme18I
    >< BspAI >< HincII           >< AvaII
    >< Bsp143I                  >< AsuI
ACACAAGATC ATGTTGACAT ATTGGGACCT CTTTCTGCTC AAACAGGAAT TGCCGTCTTA GATATGTGTG
10720      10730      10740      10750      10760      10770      10780

                                >< StyI
                                >< RsaI
                                >< EcoT14I
                                >< Eco130I
                                >< SfcI
                                >< Csp6I
    >< Fnu4HI                    >< Fnu4HI
    >< BbvI                      >< Fnu4HI
    >< BbvI                      >< AluI >< PstI
CTGCTTTGAA AGAGCTGCTG CAGAAATGGTA TGAATGGTCG TACTATCCTT GGTAGCACTA TTTAGAAGA
10790      10800      10810      10820      10830      10840      10850

                                >< StyI
                                >< EcoT14I
                                >< Eco130I
                                >< BssT1I
                                >< MboII
                                >< MaeIII>< BsaJI
TGAGTTTACA CCATTGATG TTGTTAGACA ATGCTCTGGT GTTACCTTCC AAGGTAAGTT CAAGAAAATT
10860      10870      10880      10890      10900      10910      10920

                                >< SfaNI
                                >< SduI
                                >< NspII
                                >< Tru9I
                                >< NspII
                                >< Bsp1286I
                                >< MseI
                                >< Tru9I
                                >< Bsp1286I
                                >< MseI
                                >< BmyI
                                >< FokI
                                >< RsaI ><
                                >< TfiI Csp6I ><
                                >< HinfI AfaI ><
GTTAAGGGCA CTCATCATTG GATGCTTTTA ACTTTCTTGA CATCACTATT GATTCTTGTT CAAAGTACAC
10930      10940      10950      10960      10970      10980      10990

                                >< XmnI
                                >< BsmI
                                >< BscCI
                                >< Asp700I
                                >< MunI
                                >< Fnu4HI >
                                >< BspWI ><
                                >< MaeIII
                                >< Asp700I
                                >< BbvI BbvI >
AGTGCTCACT GTTTTCTTT GTTTACGAGA ATGCTTTCTT GCCATTTACT CTTGGTATTA TGGCAATTGC
11000      11010      11020      11030      11040      11050      11060

                                >< NspI
                                >< NspHI
                                >< NlaIII
                                >< Tru9I
                                >< MseI
                                >< BsmI
                                >< BspWI >< Fnu4HI>< BspWI >< BscCI
                                >< MaeIII
TGCATGTGCT ATGCTGCTTG TTAAGCATAA GCACGCATTC TTGTGCTTGT TTCTGTTACC TTCTCTTGCA
11070      11080      11090      11100      11110      11120      11130

                                >< SfaNI
                                >< RmaI
                                >< NspI
                                >< MamI
                                >< NlaIII
                                >< HphI
                                >< NheI
                                >< BspHI
                                >< Tru9I
                                >< MaeI
                                >< BsiBI >< NlaIII
                                >< BspWI >< MseI >< AccI> < NspHI>< AluI >< BsaBI >< NlaIII
ACAGTTGCTT ACTTTAATAT GGTCTACATG CCTGCTAGCT GGGTGATGCG TATCATGACA TGGCTTGAAT
11140      11150      11160      11170      11180      11190      11200

```

FIGURA 13.25

```

                                >< Tru9I
                                >< MseI
    > < RmaI                    > < Esp4I
    > < MaeI                    >< Eco57I
                                >< AluI
    TGGCTGACAC TAGCTTGTCT GGTTATAGGC TTAAGGATTG TGTTATGTAT GCTTCAGCTT TAGTTTTGCT
      11210      11220      11230      11240      11250      11260      11270

                                >< RmaI
                                >< MaeII
                                >< MaeI
    > < NlaIII >< SfaNI >< Fnu4HI
    >< BspHI >< AluI >< BbvI >< AflIII
    TATTCTCATG ACAGCTCGCA CTGTTTATGA TGATGCTGCT AGACGTGTTT GGACACTGAT GAATGTCATT
      11280      11290      11300      11310      11320      11330      11340

                                >< Sau96I
                                >< Pali
                                >< NspIV
                                >< NlaIII
                                >< HaeIII
                                >< Sau3AI > < DdeI
                                >< NdeII >< Cfr13I
                                >< MboI >< BsuRI
                                >< DpnII >< BsiZI
                                >< DpnI >< BshI
                                >< Bsp143I > < BfrI
                                >< AccI >< BspAI>< AluI >< AsuI
    ACACTTGTTT ACAAAGTCTA CTATGGTAAT GCTTTAGATC AAGCTATTTC CATGTGGGCC TTAGTTATTT
      11350      11360      11370      11380      11390      11400      11410

                                >< RmaI
                                >< NlaIII
                                >< MaeI>< SfcI
    >< MaeIII >< MnlI >< MaeIII >< AluI>< AluI
    CTGTAACCTC TAACTATTCT GGTGTCGTTA CGACTATCAT GTTTTTAGCT AGAGCTATAG TGTTTGTGTG
      11420      11430      11440      11450      11460      11470      11480

                                DdeI >
                                >< BsrI >< NlaIII BfrI >
    TGTTGAGTAT TACCCATTGT TATTTATTAC TGGCAACACC TTACAGTGTA TCATGCTTGT TTATTGTTTC
      11490      11500      11510      11520      11530      11540      11550

                                >< Pali
                                >< HaeIII
                                >< Fnu4HI >< BsuRI
    >< BbvI >< Fnu4HI >< BspWI
    >< BbvI >< BspWI >< BshI >< Eco57I >< MaeIII
    TTAGGCTATT GTTGCTGCTG CTACTTTGGC CTTTCTGTT TACTCAACCG TTACTTCAGG CTTACTCTTG
      11560      11570      11580      11590      11600      11610      11620

                                >< ScrFI
                                >< MvaI
                                >< EcoRII
                                >< Ecl136I
                                >< DsaV
                                >< BstOI
                                >< BstNI
                                >< BsiLI
                                > < BsaJI
                                >< BsaJI
    >< Eco31I
    >< BsmAI
    >< BsaI

```

FIGURA 13. 26

```

        >< DrdI >< Alw26I
GTGTTTATGA CTACTTGGTC TCTACACAAG AATTTAGGTA TATGAACTCC CAGGGGCTTT TGCCTCCTAA
    11630      11640      11650      11660      11670      11680      11690

        >< Tru9I
        >< MseI
>< SfaNI          > < HindIII> < Tru9I
    >< MnlI          >< AluI > < MseI > < MnlI          > < NlaIII
GAGTAGTATT GATGCTTTCA AGCTTAACAT TAAGTTGTTG GGTATTGGAG GTAAACCATG TATCAAGGTT
    11700      11710      11720      11730      11740      11750      11760

        >< VneI
        >< SnaI
            >< SduI
            >< NspII
            >< HgiAI
            >< Bsp1286I
            >< BmyI >< RsaI
    >< RsaI          >< ApaLI          >< MboII
    >< Csp6I          >< Alw44I          >< Csp6I          DdeI >
    >< AfaI          >< MaeII          >< Alw21I          >< AfaI          BfrI >
GCTACTGTAC AGTCTAAAAT GTCTGACGTA AAGTGCACAT CTGTGGTACT GCTCTCGGTT CTCAACAAC
    11770      11780      11790      11800      11810      11820      11830

        >< NspII> < RsaI
            >< OraiII
            >< SduI>< Csp6I
            >< Bsp1286I
    >< MboII
        >< HinfI >< PheI          >< BmyI > < AfaI          >< MboII
TTAGAGTAGA GTCATCTTCT AAATTGTGGG CACAATGTGT ACAACTCCAC AATGATATTC TTCTTGCAAA
    11840      11850      11860      11870      11880      11890      11900

        >< TthHB8I
        >< TaqI
            >< HindIII          >< MboII          >< SfcI ><
            >< AluI          > < Eco57I          >< NlaIII
AGACACAACT GAAGCTTTCG AGAAGATGGT TTCTCTTTTG TCTGT'TTTGC TATCCATGCA GGGTGCTGTA
    11910      11920      11930      11940      11950      11960      11970

    >< VspI
    >< Tru9I          > < Ksp632I
    >< MseI          >< TthHB8I          > < EarI
    >< AsnI          >< TaqI >< MboII          > < Eam1104I
    >< AseI>< MnlI >< BcgI/a          >< Eco57I          >< Eco57I >< BcgI
GACATTAATA GGTGTGCGA GGAAATGCTC GATAACCGTG CTACTCTTCA GGCTATTGCT TCAGAATTTA
    11980      11990      12000      12010      12020      12030      12040

            >< StuI
            >< ScrFI
            >< Pali
            >< MvaI>< HaeIII
    >< EcoRII>< Eco147I
            >< Ecl136I
    >< DsaV >< BsuRI
            >< BstOI
            >< BstNI
            >< BspWI
            >< BsiLI
        >< Fnu4HI          >< BsaJI >< BshI          TfiI ><
    >< NdeI          >< BspWI>< MnlI >< BglI          >< SfcI HinfI ><
        >< AciI          >< ApyI>< AatI          > < AluI

```

FIGURA 13. 27

ES 2 396 127 T3

```

GTTCTTTACC ATCATATGCC GCTTATGCCA CTGCCCAGGA GGCCTATGAG CAGGCTGTAG CTAATGGTGA
 12050      12060      12070      12080      12090      12100      12110

      >< XmnI      >< Tru9I      >< SfaNI
      >< HphI      >< MseI      >< DdeI
      >< Asp700I   >< Eco57I   >< BbvI Fnu4HI ><
TTCTGAAGTC GTTCTCAAAA AGTTAAAGAA ATCTTTGAAT GTGGCTAAAT CTGAGTTTGA CCGTGATGCT
 12120      12130      12140      12150      12160      12170      12180

                                      XhoII ><
                                      Sau3AI ><
                                      NdeII ><
                                      MnlI >
                                      >< MnlI
                                      >< MflI
                                      >< MboI
      > < Sau3AI
      > < NdeII
      > < MboI
      > < DpnII
      >< DpnI
      >< BspWI
      > < BspAI
      >< Bsp143I
      >< Bsp143I
      >< AfaIBqlII ><
>< NlaIII
GCCATGCAAC GCAAGTTGGA AAAGATGGCA GATCAGGCTA TGACCCAAAT GTACAAACAG GCAAGATCTG
 12190      12200      12210      12220      12230      12240      12250

      >< SpeI
      >< RmaI
      >< MaeIII
      >< MaeI
      >< MboII
      >< BspWI
      >< EarI>< BfrI >< AluI
AGGACAAGAG GGCAAAAGTA ACTAGTGCTA TGCAACAAGT GCTCTTCACT ATGCTTAGGA AGCTTGATAA
 12260      12270      12280      12290      12300      12310      12320

      >< ThaI
      >< MvnI
      >< HinPII
      >< Hin6I
      >< HhaI
      >< CfoI
      >< BstUI
      >< Tru9I
      >< MseI
      >< Bsp50I
      >< AccII
      >< SfcI ><
TGATGCACTT AACAAACATTA TCAACAATGC GCGTGATGGT TGTGTTCCAC TCAACATCAT ACCATTGACT
 12330      12340      12350      12360      12370      12380      12390

      >< RsaI
      >< NlaIV
      >< Eco64I
      >< Csp6I
      >< BslI
      >< BsiYI>< KpnI
      >< BscBI
      >< BanI
      >< Asp718
      >< AfaI
      >< AccB1I
      >< Acc65I
      >< MaeIII
      >< BsgI ><
      >< Fnu4HI >< BbvI
      >< BstXI
      >< NlaIII
ACAGCAGCCA AACTCATGGT TGTTGTCCCT GATTATGGTA CCTACAAGAA CACTTGTGAT GGTAACACCT
 12400      12410      12420      12430      12440      12450      12460

      >< Zsp2I
      >< Ppu10I

```

FIGURA 13.28

```

>< NsiI
>< Mph1103I
>< NdeI>< EcoT22I
>< AvaIII >< SfaNI
>< SfaNI >< AciI
DdeI ><
BfrI ><
TTACATATGC ATCTGCACTC TGGGAAATCC AGCAAGTTGT TGATGCGGAT AGCAAGATTG TTCAACTTAG
12470 12480 12490 12500 12510 12520 12530

>< PstI
>< HaeIII >< MnlI >< DdeIDdeI ><
>< BsuRI >< MaeIII >< BspWI
>< Tru9I>< NlaIII
>< BspWI
>< MseI>< HphI >< XcmI>< BshI >< AluI BspWI ><
TGAAATTAAC ATGGACAATT CACCAAATTT GGCTTGGCCT CTTATTGTTA CAGCTCTAAG AGCCAACTCA
12540 12550 12560 12570 12580 12590 12600

RsaI ><
NlaIV ><
KpnI ><
>< Fnu4HI
Eco64I ><
Csp6I ><
BscBI ><
Asp718 ><
AfaI ><
>< AciI>< Bani
AccB1I ><
>< MseI >< HinfI >< PstI
>< AluI >< SfcI >< DdeI>< BsrI >< PshAI Acc65I ><
GCTGTTAAAC TACAGAATAA TGAAGTGGT CCAGTAGCAC TACGACAGAT GTCCTGTGCG GCTGGTACCA
12610 12620 12630 12640 12650 12660 12670

>< TthHB8I
>< TaqI
>< SfuI
>< NspV
>< MnlI
>< LspI
>< Csp45I
>< BstBI
>< Bsp119I
>< BsiCI
>< Bpu14I
>< AsuII
>< RsaI
>< Csp6I
>< AluI
>< AfaI
CACAAACAGC TTGTACTGAT GACAATGCAC TTGCCTACTA TAACAATTTCG AAGGGAGGTA GGTGTGTGCT
12680 12690 12700 12710 12720 12730 12740

>< XhoII
>< Sau3AI
>< NdeII
>< MflI
>< MboI
>< DpnII
>< DpnI
>< BstYI
>< BspAI
>< Bsp143I
>< BglII
>< TfiI
>< RmaI
>< HinfI
>< MaeI
>< DdeI
>< RsaI
>< Csp6I
>< Csp6I>< RsaI
>< AfaI>< AfaI
GGCATTACTA TCAGACCACC AAGATCTCAA ATGGGCTAGA TTCCCTAAGA GTGATGGTAC AGGTACAATT
12750 12760 12770 12780 12790 12800 12810

>< Sau96I
>< PssI
>< PstI
>< NspIV

```

FIGURA 13.29

```

                                >< HaeIII
                                >< EcoO109I
                                >< DraII
                                >< Cfr13I
                                >< BsuRI
                                >< BsiZI      RsaI >
                                >< BshI      Csp6I ><
                                >< AsuI      AfaI >
TACACAGAAC TGGAACCACC TTGTAGGTTT GTTACAGACA CACCAAAAGG GCCTAAAGTG AAATACTTGT
12820      12830      12840      12850      12860      12870      12880

                                >< SfcI
                                > < MboII
                                MaeII ><
                                >< Fnu4HI >< RsaI
                                >< Eco57I >< Csp6I
                                > < BbsI
                                >< Tru9I
                                >< MseI >< MnlI
                                >< BbvI      >< AluI      >< AfaI
ACTTCATCAA AGGCTTAAAC AACCTAAATA GAGGTATGGT GCTGGGCAGT TTAGCTGCTA CAGTACGTCT
12890      12900      12910      12920      12930      12940      12950

                                >< RsaI
                                >< SfcI >< Csp6I
                                >< BspWI      >< AfaI      >< BspMI      AccI ><
TCAGGCTGGA AATGCTACAG AAGTACCTGC CAATTCAACT GTGCTTTCCT TCTGTGCTTT TGCAGTAGAC
12960      12970      12980      12990      13000      13010      13020

                                >< RmaI
                                >< MnlI
                                >< MaeI      >< HphI
CCTGCTAAAG CATATAAGGA TTACCTAGCA AGTGGAGGAC AACCAATCAC CAACTGTGTG AAGATGTTGT
13030      13040      13050      13060      13070      13080      13090

                                >< SinI
                                >< Sau96I
                                >< NspIV
                                >< NspHII
                                >< NlaIII
                                >< Eco47I
                                >< Eam1105I
                                >< Cfr13I
                                >< BsiZI
                                >< Bme18I >< XcmI
                                >< AvaII   PleI ><
                                >< AfaI      >< AfaI      >< MaeIII      >< AluI      >< AsuI> < HinfI
GTACACACAC TGGTACAGGA CAGGCAATTA CTGTAACACC AGAAGCTAAC ATGGACCAAG AGTCCTTTGG
13100      13110      13120      13130      13140      13150      13160

                                >< TfiI
                                >< MaeIII
                                >< SfaNI
                                >< NlaIII      >< FokI
                                >< HinfI
TGGTGCTTCA TGTGTCTGT ATTGTAGATG CCACATTGAC CATCCAAATC CTAAAGGATT CTGTGACTTG
13170      13180      13190      13200      13210      13220      13230

                                > < RsaI
                                >< MaeII
                                >< Csp6I
                                > < AfaI
                                >< DdeI
                                >< BfrI
AAAGGTAAGT ACGTCCAAAT ACCTACCACT TGTGCTAATG ACCCAGTGGG TTTTACACTT AGAAACACAG
13240      13250      13260      13270      13280      13290      13300

                                >< ThaI

```

FIGURA 13. 30

```

                                >< SfaNI
                                >< MvnI
                                >< BstUI
                                >< Bsp50I
                                >< AciI
>< RsaI
>< Csp6I
>< AfaI >< AciI                                >< SfcI >< MaeIII                                >< AccIISfaNI ><
TCTGTACCGT CTGCGGAATG TGGAAAGGTT ATGGCTGTAG TTGTGACCAA CTCCGCGAAC CCTTGATGCA
13310      13320      13330      13340      13350      13360      13370

                                >< Zsp2I
                                > < SfaNI
                                >< Mph1103I>< Tru9I
>< Ppu10I>< MaeII                                Fnu4HI ><
                                >< NsiI> < FokI                                BsgI ><
                                >< EcoT22I >< MseI                                >< BbvI
>< AciI>< AvaIII                                >< DraI                                >< AciI                                >< Fnu4HI                                AciI ><
GTCTGCGGAT GCATCAACGT TTTTAAACGG GTTTGCGGTG TAAGTGCAGC CCGTCTTACA CCGTGCGGCA
13380      13390      13400      13410      13420      13430      13440

>< SpeI
>< ScaI
>< RsaI
>< RmaI
>< MaeI
> < Csp6I                                >< SfcI                                >< BspWI
>< BspWI >< AfaI                                >< AccI                                >< BcgI/a                                BcgI >
CAGGCACTAG TACTGATGTC GTCTACAGGG CTTTGTATAT TTACAACGAA AAAGTTGCTG GTTTTGCAAA
13450      13460      13470      13480      13490      13500      13510

                                >< ScrFI
                                >< MvaI
                                >< MnlI
>< EcoRII
>< Ecl136I
>< BstOI
>< BstNI
>< BslI
>< DsaV >< BsiYI
>< BsiLI
>< ApyI                                >< PleI
                                > < FokI >< HinfI
GTTTCCTAAAA ACTAATTGCT GTCGCTTCCA GGAGAAGGAT GAGGAAGGCA ATTTATTAGA CTCTTACTTT
13520      13530      13540      13550      13560      13570      13580

                                >< NlaIII
                                >< Ksp632I
                                >< EarI
                                >< Eam1104I
>< Tru9I
>< MseI                                >< BsmAI
>< MnlI                                >< Alw26I                                >< MboII                                >< Tru9I
                                >< MseI
GTAGTTAAGA GGCATACTAT GTCTAACTAC CAACATGAAG AGACTATTTA TAACTTGGTT AAAGATTGTC
13590      13600      13610      13620      13630      13640      13650

                                >< RsaI
                                >< NlaIV
> < NlaIII
                                >< KpnI
                                >< HphI
> < Eco64I
>< Csp6I
>< BscBI
> < BanI
> < Asp718

```

FIGURA 13. 31


```

                                >< MaeIII >< AfaI
                                > < AccBII MaeII ><
>< NspBII                                >< Acc65I > < HgaI
>< AciI                                >< NlaIII
CAGCGGTTGC TGTCCATGAC TTTTCAAGT TTAGAGTAGA TGGTGACATG GTACCACATA TATCACGTCA
13660      13670      13680      13690      13700      13710      13720

                                >< MnlI
                                >< MaeII
GCGTCTAACT AAATACACAA TGGCTGATTT AGTCTATGCT CTACGTCATT TTGATGAGGG TAATTGTGAT
13730      13740      13750      13760      13770      13780      13790

>< Tru9I
>< MseI                                >< MaeIII >< MnlI
ACATTAAAAG AAATACTCGT CACATACAAT TGCTGTGATG ATGATTATTT CAATAAGAAG GATTGGTATG
13800      13810      13820      13830      13840      13850      13860

                                >< ThaI
                                >< MvnI
                                >< MluI
                                >< BstUI
                                >< Bsp50I
                                >< RsaI
                                >< HphI
>< TfiI                                >< AflIII >< DdeI >< Csp6I Tru9I ><
>< HinfI                                >< AccII >< BfrI >< AfaI MseI ><
ACTTCGTAGA GAATCCTGAC ATCTTACGCG TATATGCTAA CTTAGGTGAG CGTGTACGCC AATCATTATT
13870      13880      13890      13900      13910      13920      13930

                                XhoII >
                                Sau3AI >
                                NdeII >
                                MflI >
                                MboI >
> < SfaNI                                >< RsaI                                MboI >
>< RsaI                                > < Csp6I                                DpnII >
>< Csp6I                                >< BspWI                                BstYI >
>< AfaI                                >< SfaNI                                >< AfaI                                BspAI >
AAAGACTGTA CAATTCTGCG ATGCTATGCG TGATGCAGGC ATTGTAGGCG TACTGACATT AGATAATCAG
13940      13950      13960      13970      13980      13990      14000

                                > < ScrFI
                                > < MvaI
                                >< Fnu4HI
>< Tru9I                                >< EcoRII
>< MseI                                > < Ecl136I
>< DpnI                                > < BstOI
>< Bsp143I                                > < BstNI
>< AlwI                                >< RsaI                                >< BslI
>< AfaI                                > < HphI                                >< BsiYI
>< Bsp6I                                >< Csp6I                                > < BsiII
>< BbvI                                > < BbvI > < ApyI
>< AfaI                                >< AfaI                                >< DsaV >< AciI
GATCTTAATG GGAAGTGGTA CGATTTCGGT GATTTCGTAC AAGTAGCACC AGGCTGCGGA GTTCCTATTG
14010      14020      14030      14040      14050      14060      14070

                                >< SfaNI
                                >< RmaI                                > < HinfI
>< TfiI                                >< MnlI                                >< Fnu4HIIpleI ><
>< HinfI                                >< BsiBI                                >< MaeI                                >< DdeI
>< FokI                                >< BsaBI                                >< BbvI                                >< BspWI NdeI ><
TGGATTCATA TTAATCATTG CTGATGCCCA TCCTCACTTT GACTAGGGCA TTGGCTGCTG AGTCCCATAT
14080      14090      14100      14110      14120      14130      14140

>< Sau3AI
>< NdeII

```

FIGURA 13. 32

```

>< MboI
>< MamI
>< DpnII
    >< DpnI
        >< BspWI
            >< BspAI
                >< BspI43I
                    >< XcmI
                        >< BsiBI
                            >< Tru9I
                                >< BsaBI >< FokI
                                    >< MseI
                                        GGATGCTGAT CTCGCAAAAC CACTTATTAA GTGGGATTTG CTGAAATATG ATTTTACGGA AGAGAGACTT
                                            14150      14160      14170      14180      14190      14200      14210

                                                > < SinI
                                                    > < Sau96I
                                                        > < NspIV
                                                            >< NspHII
                                                                >< NlaIV
                                                                    >< TthHB8I
                                                                        >< TaqI
                                                                            >< McrI
                                                                                > < Eco47I
                                                                                    > < Ksp632I
                                                                                        > < CfrI3I
                                                                                            > < EarI
                                                                                                > < Bsi2I
                                                                                                    > < Eam1104I
                                                                                                        >< SspI>< BscBI
                                                                                                            >< BsmAI
                                                                                                                > < Tru9I
                                                                                                                    > < Bme18I
>< MboII >< BsiEI> < MseI > < AvaII
    >< Alw26I >< DraI > < AsuI
        TGTCTCTTCG ACCGTTATTT TAAATATTGG GACCAGACAT ACCATCCCAA TTGTATTAACT TGTTTGGATG
            14220      14230      14240      14250      14260      14270      14280

                                                                SinI ><
                                                                Sau96I ><
                                                                NspIV ><
                                                                NspHII >
                                                                Eco47I ><
                                                                CfrI3I ><
                                                                Bsi2I ><
                                                                Bme18I ><
                                                                AvaII ><
                                                                AsuI ><
                                                                >< Tru9I
                                                                >< MseI
>< FokI
    ATAGGTGTAT CCTTCATTGT GCAAACCTTA ATGTGTTATT TTCTACTGTG TTTCCACCTA CAAGTTTTGG
        14290      14300      14310      14320      14330      14340      14350

>< SpeI
>< RmaI
>< MaeI
    >< SspI
        ACCACTAGTA AGAAAAATAT TTGTAGATGG GTTTCCTTTT GTTGTTTCAA CTGGATACCA TTTTCGTGAG
            14360      14370      14380      14390      14400      14410      14420

                                                                >< ThaI>< Esp3I
                                                                >< DdeI
                                                                >< BstUI
                                                                >< Bsp50I >< BsmBI
                                                                >< MvnI>< BsmAI
                                                                >< HgaI>< AluI >< Alw26I
                                                                >< FokI >< AccII
                                                                > < BbvI
TTAGGAGTCG TACATAATCA GGATGTAAAC TTACATAGCT CGCGTCTCAG TTTCAAGGAA CTTTGTAGTGT
    14430      14440      14450      14460      14470      14480      14490

                >< Zsp2I
                >< SphI
>< Ppu10I
    >< PaeI
    >< NspI

```

FIGURA 13. 33

```

>< Sau3AI      >< NspHI
>< NdeII       >< NsiI
>< MboI        >< NlaIII
>< DpnII       >< Mph1103I
> < DpnI      >< Fnu4HI
>< Fnu4HI>< BspWI >< EcoT22I
>< BspAI       >< BspWI
> < Bsp143I> < AvaIII > < AlwNI
>< AlwI        >< AluI    >< AluI    >< BbvI    >< MaeI
ATGCTGCTGA TCCAGCTATG CATGCAGCTT CTGGCAATTT ATTGCTAGAT AAACGCACTA CATGCTTTTC
14500      14510      14520      14530      14540      14550      14560

>< ScrFI
>< NciI
>< MspI
>< HpaII
>< HapII
>< Fnu4HI
>< AlwNI
>< AluI
AGTAGCTGCA CTAACAAACA ATGTTGCTTT TCAAAGTGTG AAACCCGGTA ATTTTAATAA AGACTTTTAT
14570      14580      14590      14600      14610      14620      14630

>< Tru9I
>< MseI
GACTTTGCTG TGTCTAAAGG TTTCTTTAAG GAAGGAAGTT CTGTTGAAGT AAAACACTTC TTCTTTGCTC
14640      14650      14660      14670      14680      14690      14700

>< FokI
>< Fnu4HI
AGGATGGCAA CGCTGCTATC AGTGATTATG ACTATTATCG TTATAATCTG CCAACAATGT GTGATATCAG
14710      14720      14730      14740      14750      14760      14770

>< VspI
>< Tru9I
>< MseI
>< AsnI
>< AseI
>< MaeIII
ACAACCTCTA TTCGTAGTTG AAGTTGTTGA TAAATACTTT GATTGTTACG ATGGTGGCTG TATTAATGCC
14780      14790      14800      14810      14820      14830      14840

>< Tru9I
>< MseI
>< HpaI
>< HindII
>< HincII
AACCAAGTAA TCGTTAACAA TCTGGATAAA TCAGCTGGTT TCCCATTTAA TAAATGGGGT AAGGCTAGAC
14850      14860      14870      14880      14890      14900      14910

>< SfaNI
>< Sau3AI
>< NdeII
>< MboI
>< DpnII
>< DpnI
>< Bsp143I
>< PleI
>< HinfI>< MnlI
>< BspAI >< AlwI
TTTATTATGA CTCAATGAGT TATGAGGATC AAGATGCACT TTTCGCGTAT ACTAAGCGTA ATGTCATCCC
14920      14930      14940      14950      14960      14970      14980

>< SstI
>< SduI
>< SacI

```

FIGURA 13. 34

```

                                >< NspII
                                >< HgiAI
                                >< Eco24I
                                > < Ecl136II
                                >< Bsp1286I
                                >< BmyI
                                >< BanII
                                >< Alw21I
                                > < AluI
                                >< AluI
TACTATAACT CAAATGAATC TTAAGTATGC CATTAGTGCA AAGAATAGAG CTCGCACCGT AGCTGGTGTCT
14990      15000      15010      15020      15030      15040      15050

                                >< Tru9I
                                >< TfiI
                                >< MseI
                                >< HinfI
                                > < Esp4I
                                > < AflIII
                                >< BspWI
                                > < AluI
                                >< AluI
TACTATAACT CAAATGAATC TTAAGTATGC CATTAGTGCA AAGAATAGAG CTCGCACCGT AGCTGGTGTCT
14990      15000      15010      15020      15030      15040      15050

                                >< ScaI
                                >< SfcI>< RsaI
                                >< BsmAI >< Csp6I
                                >< Alw26I >< AfaI
TCTATCTGTA GTACTATGAC AAATAGACAG TTTCATCAGA AATTATTGAA GTCAATAGCC GCCACTAGAG
15060      15070      15080      15090      15100      15110      15120

                                >< Tru9I
                                >< MseI
GAGCTACTGT GGTAATTGGA ACAAGCAAGT TTTACGGTGG CTGGCATAAT ATGTTAAAAA CTGTTTACAG
15130      15140      15150      15160      15170      15180      15190

                                NspI ><
                                NspHI ><
                                NlaIII ><
                                >< NlaIII
                                DdeI ><
                                BspWI ><
                                >< MaeIII
                                BfrI ><
TGATGTAGAA ACTCCACACC TTATGGGTTG GGATTATCCA AAATGTGACA GAGCCATGCC TAACATGCTT
15200      15210      15220      15230      15240      15250      15260

                                > < Pali
                                > < HaeIII
                                >< BsuRI
                                > < BshI
                                >< MnlI
                                >< MaeIII
                                SfcI ><
AGGATAATGG CCTCTCTTGT TCTTGCTCGC AAACATAACA CTTGCTGTAA CTTATCACAC CGTTTCTACA
15270      15280      15290      15300      15310      15320      15330

                                Tru9I ><
                                ScrFI >
                                MvaI >
                                >< MseI
                                FokI ><
                                EcoRII ><
                                Ecl136I >
                                DsaV ><
                                BstOI >
                                BstNI >
                                >< NlaIII
                                > < Fnu4HI
                                BsiLI >
                                >< AluI
                                >< AviIII >< MseI
                                >< AciI
                                ApyI >
GGTTAGCTAA CGAGTGTGCG CAAGTATTAA GTGAGATGGT CATGTGTGGC GGCTCACTAT ATGTTAAACC
15340      15350      15360      15370      15380      15390      15400

                                > < SfaNI
                                >< MspI
                                >< HpaII
                                >< HphI
                                >< HapII
                                >< BspWI
                                >< Tru9I
                                >< MaeIII ><
                                >< MseI
                                >< AluI ><

```

FIGURA 13. 35

ES 2 396 127 T3

```

AGGTGGAACA TCATCCGGTG ATGCTACAAC TGCTTATGCT AATAGTGTCT TTAACATTG TCAAGCTGTT
15410      15420      15430      15440      15450      15460      15470

>< BspWI
ACAGCCAATG TAAATGCACT TCTTTCAACT GATGGTAATA AGATAGCTGA CAAGTATGTC CGCAATCTAC
15480      15490      15500      15510      15520      15530      15540

>< DrdI
>< AluI
> < AciI

>< Sau3AI
>< NdeII
>< MboI
> < MamI
>< FbaI
>< DpnII
>< DpnI
>< BspHI
>< BspAI
>< BspI43I
>< BsiQI
>< SfcI
>< BsmAI
>< Alw26I
>< BsiBI>< NlaIII
>< BsaBI>< FokI
>< BclI>< EcoRI
AACACAGGCT CTATGAGTGT CTCTATAGAA ATAGGGATGT TGATCATGAA TTCGTGGATG AGTTTTACGC
15550      15560      15570      15580      15590      15600      15610

>< TfiI
>< SfaNI
>< NlaIII
>< BspMI
>< HinfI
>< MaeIII
TTACCTGCGT AACATTTCT CCATGATGAT TCTTTCTGAT GATGCCGTTG TGTGCTATAA CAGTAACTAT
15620      15630      15640      15650      15660      15670      15680

> < RmaI
>< NheI >< Tru9I
>< Fnu4HI
> < MaeI
>< Tru9I
>< AciI
>< AluI >< MseI
>< MseI
MnlI ><
GCGGCTCAAG GTTTAGTAGC TAGCATTAAG AACTTTAAGG CAGTTCTTTA TTATCAAAAT AATGTGTTCA
15690      15700      15710      15720      15730      15740      15750

>< SinI
>< Sau96I
>< PssI
>< Psp5II
>< PpuMI
>< NspIV
>< NspHII
>< EcoO109I
>< Eco47I
>< DraII
>< Cfr13I
>< BsiZI
>< Bmel8I
>< DdeI
>< NlaIII
>< BsmAI
>< DdeI
>< Alw26I
>< AsuI
>< MnlI
TGTCTGAGGC AAAATGTTGG ACTGAGACTG ACCTTACTAA AGGACCTCAC GAATTTTGCT CACAGCATAC
15760      15770      15780      15790      15800      15810      15820

>< XhoII
>< Sau3AI
>< NdeII
>< MflI
>< MboI

```

FIGURA 13. 36

```

                >< RsaI          >< DpnII
                >< MaeII        >< DpnI
                >< Csp6I        >< BstYI
                >< BsaAI        >< BspMI
                >< AflIII       >< BspAI
>< BspWI>< MseI          >< AfaI  >< AlwI>< Bsp143I
AATGCTAGTT AAACAAGGAG ATGATTACGT GTACCTGCCT TACCCAGATC CATCAAGAAT ATTAGGCGCA
15830      15840      15850      15860      15870      15880      15890

                >< RsaI          >< SfaNI
                >< Csp6I        >< MaeIII
                >< AfaI        >< BsrI ><
GGCTGTTTTG TCGATGATAT TGTCAAAACA GATGGTACAC TTATGATTGA AAGGTTCTGT TCACTGGCTA
15900      15910      15920      15930      15940      15950      15960

> < FokI
>< BspWI
TTGATGCTTA CCCACTTACA AAACATCCTA ATCAGGAGTA TGCTGATGTC TTTCACCTGT ATTTACAATA
15970      15980      15990      16000      16010      16020      16030

                >< Van91I
                >< PflMI
                >< NspI
                > < Pali>< NspHI
                > < MscI>< NlaIII
                > < HaeIII
                > < BsuRI
                >< BsrI
                >< EaeI >< BslI >< NspI
                > < BshI>< BsiYI >< NspHI
                >< NlaIII >< AflIII >< AflIII
                >< MaeIII >< AluI > < BalI>< AccB7I >< NlaIII
CATTAGAAAG TTACATGATG AGCTTACTGG CCACATGTTG GACATGTATT CCGTAATGCT AACTAATGAT
16040      16050      16060      16070      16080      16090      16100

                >< RsaI> < NlaIV
                >< MnlI
                >< Csp6I >< DdeI
                >< BsrI >< MnlI
                >< AfaI> < BscBI
AACACCTCAC GGTACTGGGA ACCTGAGTTT TATGAGGCTA TGTCACACC ACATACAGTC TTGCAGGCTG
16110      16120      16130      16140      16150      16160      16170

                >< NlaIV
                >< EcoNI
                >< Eco31I
                >< Eco64I>< BsmAI
                >< BscBI >< BslI
                >< BanI >< BsiYI
                >< AciI >< BsaI
                >< AccB1I>< Alw26I BbvI ><
>< BspWI
TAGGTGCTTG TGTATTGTGC AATTCACAGA CTTCACTTCG TTGCGGTGCC TGTATTAGGA GACCATTCTT
16180      16190      16200      16210      16220      16230      16240

                >< Tth111I
                >< Fnu4HI >< NlaIII
                >< BspWI >< AspI
                >< Tru9I
                >< MseI
ATGTTGCAAG TGCTGCTATG ACCATGTCAT TTCAACATCA CACAAATTAG TGTTGTCTGT TAATCCCTAT
16250      16260      16270      16280      16290      16300      16310

                >< ScrFI
                >< MvaI

```

FIGURA 13. 37

```

>< EcoRII
>< Ecl136I
>< DsaV
>< BstOI
>< BstNI
>< BsiLI
>< BsaJI
>< ApyI
>< MaeIII >< MaeIII
>< MaeI
>< RmaI
>< MnlI
BspWI ><
>< AluI
GTTTGCAATG CCCAGGTTG TGATGTCAC TATGTGACAC AACTGTATCT AGGAGGTATG AGCTATTATT
16320 16330 16340 16350 16360 16370 16380

>< MaeIII >< MnlI
GCAAGTCACA TAAGCCTCCC ATTAGTTTTC CATTATGTGC TAATGGTCAG GTTTTTGGTT TATACAAAAA
16390 16400 16410 16420 16430 16440 16450

>< NspI >< NspI
>< NspHI >< Tth111I >< NspHI
>< NlaIII>< MaeIII>< MaeIII >< NlaIII
>< AflIII >< AspI >< AflIII
CACATGTGTA GGCAGTGACA ATGTCAC TTTCAATGCG ATAGCAACAT GTGATTGGAC TAATGCTGGC
16460 16470 16480 16490 16500 16510 16520

>< RsaI
>< P1eI
>< DdeI
>< Csp6I
>< BsmAI >< HinfI >< MnlI
>< Alw26I >< HindIII DdeI ><
>< AfaI >< AluI >< Fnu4HI >< BbvI
GATTACATAC TTGCCAACAC TTGTACTGAG AGACTCAAGC TTTTCGCAGC AGAAACGCTC AAAGCCACTG
16530 16540 16550 16560 16570 16580 16590

>< ThaI
>< ScaI
>< RsaI >< RsaI
>< MvnI
>< Csp6I >< Csp6I
>< BstUI
>< Bsp50I
>< Tru9I
>< MseI >< NdeI >< AfaI >< AfaI
>< AluI >< AccII MnlI >
AGGAAACATT TAAGCTGTCA TATGGTATTG CCACTGTACG CGAAGTACTC TCTGACAGAG AATTGCATCT
16600 16610 16620 16630 16640 16650 16660

MaeIII ><
>< MaeIII
>< Eco065I
>< Eco91I
>< BstPI
>< BstEII
>< BsrI
>< SfaNI >< RmaI
>< NlaIII >< MaeI
TTCATGGGAG GTTGGAACAC CTAGACCACC ATTGAACAGA AACTATGTCT TTACTGGTTA CCGTGTAAC
16670 16680 16690 16700 16710 16720 16730

RsaI ><
>< MnlI
>< HphI
>< RsaI >< RsaI
>< Csp6I >< Csp6I >< SfaNI Csp6I ><
>< AfaI >< AfaI >< MaeIII >< HphI AfaI ><
AAAAATAGTA AAGTACAGAT TGGAGAGTAC ACCTTTGAAA AAGGTGACTA TGGTGATGCT GTTGTGTACA
16740 16750 16760 16770 16780 16790 16800

```

FIGURA 13. 38

```

    >< RsaI
    >< Csp6I
    >< AfaI
GAGGTACTAC GACATACAAG TTGAATGTTG GTGATTACTT TGTGTTGACA TCTCACACTG TAATGCCACT
16810      16820      16830      16840      16850      16860      16870

    >< HphI
    >< HindII
    >< HincII
    DdeI ><
    BfrI ><

    >< VneI
    >< SnaI
    >< SduI
    >< NspII
    >< HgiAI
    > < SduI
    > < NspII
    >< DraIII
    >< Bsp1286I
    >< BmyI
    >< BspWI
    >< DraIII
    >< RsaI
    >< Csp6I
    >< ApaLI
    >< RmaI
    >< Bsp1286I
    >< Csp6I
    >< Alw44I
    >< MaeI
    >< BmyI
    >< BsrI
    >< Alw21I
    >< AfaI
    DdeI >
TAGTGCACCT ACTCTAGTGC CACAAGAGCA CTATGTGAGA ATTACTGGCT TGTACCCAAC ACTCAACATC
16880      16890      16900      16910      16920      16930      16940

    StyI ><
    SinI >
    Sau96I >
    NspIV >
    EcoT14I ><
    Eco47I >
    Eco130I ><
    >< ScaI
    Cfr13I >
    BssT1I ><
    >< SphI
    >< RsaI
    Bsi2I >
    >< PaeI
    BsaJI ><
    >< NlaIII
    Bme18I >
    >< NspI
    >< Csp6I
    AvaII >
    >< NspHI
    >< AfaI
    AsuI >
    >< RmaI
    >< MaeI
TCAGATGAGT TTTCTAGCAA TGTGCAAAAT TATCAAAAAGG TCGGCATGCA AAAGTACTCT ACACTCCAAG
16950      16960      16970      16980      16990      17000      17010

    >< ScrFI
    >< RsaI
    >< MvaI
    >< EcoRII
    >< Ecl136I
    > < Csp6I
    >< BstOI
    >< BstNI
    >< XcmI
    >< BslI
    >< NspHII
    >< BsiYI
    >< BsiLI
    >< ApyI
    >< BsrI
    >< DsaV
    >< AfaI
    > < HinfI
    >< P1eI
GACCACCTGG TACTGGTAAG AGTCATTTTG CCATCGGACT TGCTCTCTAT TACCCATCTG CTCGCATAGT
17020      17030      17040      17050      17060      17070      17080

    >< SfaNI
    >< SphI
    >< PvuII
    >< PaeI
    >< Psp5I
    >< NspI
    >< NspBII
    >< NspHI
    >< Fnu4HI
    > < Tru9I
    >< Bst1107I
    > < NlaIII
    >< BspWI
    >< SspI
    >< AccI
    >< NlaIII
    >< AluI
    >< BbvI
    > < MseI
GTATACGGCA TGCTCTCATG CAGCTGTTGA TGCCCTATGT GAAAAGGCAT TAAAATATTT GCCCATAGAT
17090      17100      17110      17120      17130      17140      17150

```

FIGURA 13. 39


```

> < ThaI
>< ThaI
> < MvnI
>< MvnI >< ThaI
> < HinPII
>< HinPII
>< HinPII >< MvnI
> < Hin6I
>< Hin6I
> < HhaI
>< HhaI >< HhaI
> < CfoI
>< CfoI >< CfoI
> < BstUI
>< BstUI >< BstUI
>< BssHII
>< BspMI
> < Bsp50I
>< Bsp50I>< Bsp50I
>< TfiI >< Hin6I> < AccII RmaI >
>< HinFI >< AccII >< AccII MaeI >
> < EcoRI
AAATGTAGTA GAATCATACC TGCGCGTGCG CGCGTAGAGT GTTTTGATAA ATTCAAAGTG AATTCAACAC
17160 17170 17180 17190 17200 17210 17220

>< Zsp2I
>< Ppu10I
>< NsiI
>< Mph1103I
>< EcoT22I
>< BsqI > < AvaIII >< DrdI
TAGAACAGTA TGTTTTCTGC ACTGTAAATG CATTGCCAGA AACAACTGCT GACATTGTAG TCTTTGATGA
17230 17240 17250 17260 17270 17280 17290

>< RmaI
>< MaeI >< MaeII
AATCTCTATG GCTACTAATT ATGACTTGAG TGTGTCAAT GCTAGACTTC GTGCAAAACA CTACGTCTAT
17300 17310 17320 17330 17340 17350 17360

>< Sau3AI
>< NdeII
>< MboI
>< DpnII
>< DpnI
>< BspAI >< RmaI
>< AlwI>< Bsp143I > < AciI >< MaeI SspI ><
ATTGGCGATC CTGCTCAATT ACCAGCCCCC CGCACATTGC TGACTAAAGG CACACTAGAA CCAGAATATT
17370 17380 17390 17400 17410 17420 17430

>< SinI
>< Sau96I
>< NspIV >< StyI
>< NspHII >< NspI
>< Eco47I >< NspHI
>< Cfr13I >< NlaIII
>< BsiZI >< EcoT14I
>< BsgI >< Eco130I
>< Bme18I >< BssT1I
>< AvaII >< BsaJI
>< Tru9I
>< MseI >< AsuI> < AflIII
TTAATTCAGT GTGCAGACTT ATGAAAACAA TAGGTCCAGA CATGTTCTT GGAAGTTGTC GCCGTTGTCC
17440 17450 17460 17470 17480 17490 17500

```

FIGURA 13. 40

ES 2 396 127 T3

```

                >< HindII
                >< HincII
                >< AluI
TGCTGAAATT GTTGACACTG TGAGTGCTTT AGTTTATGAC AATAAGCTAA AAGCACACAA GGATAAGTCA
17510      17520      17530      17540      17550      17560      17570

>< AluI                >< NlaIII
GCTCAATGCT TCAAAATGTT CTACAAAGGT GTTATTACAC ATGATGTTTC ATCTGCAATC AACAGACCTC
17580      17590      17600      17610      17620      17630      17640

                >< MnlI
>< EcoNI
                >< BsiI
                >< BsiYI
                >< HphI
                >< AluI
AAATAGGCGT TGTAAGAGAA TTTCTTACAC GCAATCCTGC TTGGAGAAAA GCTGTTTTTA TCTCACCTTA
17650      17660      17670      17680      17690      17700      17710

                >< SfcI                >< DdeI                >< TfiI
                > < AluI                >< BfrI                >< HinfI
TAATTCACAG AAGCTGTAG CTTCAAAAAT CTTAGGATTG CCTACGCAGA CTGTTGATTC ATCACAGGGT
17720      17730      17740      17750      17760      17770      17780

                >< Tth111I
                >< AspI
                > < HindII
                > < HincII
                >< AciI
TCTGAATATG ACTATGTCAT ATTACACAA ACTACTGAAA CAGCACACTC TTGTAATGTC AACCGCTTCA
17790      17800      17810      17820      17830      17840      17850

                >< XhoII
                >< Sau3AI
                >< NdeII
                >< MflI
                >< MboI
                >< MamI
                >< DpnII
                >< DpnI
                >< BstYI
                >< BspAI
                >< Bsp143I
                >< BsiBI
                >< BsaBI
                >< BglII
                >< BspWI
ATGTGGCTAT CACAAGGGCA AAAATTGGCA TTTTGTGCAT AATGTCTGAT AGAGATCTTT ATGACAACT
17860      17870      17880      17890      17900      17910      17920

                >< XbaI
                >< RmaI
                >< MaeI                >< MaeII
                >< MaeIII
                BsrI ><
GCAATTTACA AGTCTAGAAA TACCACGTCG CAATGTGGCT ACATTACAAG CAGAAAATGT AACTGGACTT
17930      17940      17950      17960      17970      17980      17990

                >< Sau3AI
                >< NdeII
                >< MboII
                >< MboI
                > < FokI
                >< DpnII
                >< DpnI
                >< BspAI
                >< Bsp143I
                >< Tru9I
                >< MseI>< SfcI
                >< BbsI > < BsrI
                >< NlaIV
                >< Eco64I
                >< BscBI
                >< BanI
                >< AccBI
                MnlI ><
                >< DdeI

```

FIGURA 13. 41

```

TTTAAGGACT GTAGTAAGAT CATTACTGGT CTTTCATCCTA CACAGGCACC TACACACCTC AGCGTTGATA
18000      18010      18020      18030      18040      18050      18060

      >< ScrFI
      >< MvaI
      >< EcoRII
      >< Eco57I
      >< Ecl136I
      >< DsaV
      >< BstOI
      >< BstNI
      >< HindII>< BsiLI
      >< HincII>< ApyI
      >< P1eI
      >< NlaIII
      >< HinfI ><
      >< AccI ><
TAAAGTTCAA GACTGAAGGA TTATGTGTTG ACATACCAGG CATAACCAAG GACATGACCT ACCGTAGACT
18070      18080      18090      18100      18110      18120      18130

      >< MaeIII
      >< Eco0651
      >< Eco91I
      >< BstXI
      >< BstPI
      >< BstEII
      >< HphI
      >< AccII ><
      >< ThaI ><
      >< MvnI ><
      >< BstUI ><
      >< Bsp50I ><
      >< AciI
CATCTCTATG ATGGGTTTCA AAATGAATTA CCAAGTCAAT GGTTACCCTA ATATGTTTAT CACCCGCGAA
18140      18150      18160      18170      18180      18190      18200

      >< XmnI
      > < MboII
      > < MaeIII
      >< Asp700I
      >< AluI
      >< MaeII
      >< MnlI
      >< MaeI
      >< SfaNI
      >< RmaI
      >< NlaIII
      >< MaeI
GAAGCTATTC GTCACGTTTCG TCGGTGGATT GGCTTTGATG TAGAGGGCTG TCATGCAACT AGAGATGCTG
18210      18220      18230      18240      18250      18260      18270

      >< Tru9I
      >< MseI
      >< HpaI
      >< RsaI
      >< GsuI
      >< Csp6I
      >< BpmI
      >< AfaI
      >< RmaI
      >< MnlI
      >< MaeI
      >< AluI
      >< SfcI
      >< BfrI
      >< AfaI
      >< DdeI >< AluI
      >< BsrI ><
TGGGTACTAA CCTACCTCTC CAGCTAGGAT TTTCTACAGG TGTTAACTTA GTAGCTGTAC CGACTGGTTA
18280      18290      18300      18310      18320      18330      18340

      >< ScrFI
      >< MvaI
      >< MnlI
      >< MaeIII
      >< EcoRII
      >< Eco0651
      >< EcoNI
      >< Eco91I
      >< Ecl136I
      >< DsaV
      >< Tru9I ><
      >< DraIII
      >< BstPI
      >< BstOI
      >< BstNI
      >< PmeI ><
      >< BstEII
      >< BslI
      >< MseI ><
      >< BsiYI
      >< HphI ><
      >< BsiLI
      >< DraI ><
      >< ApyI >< BsrI
      >< HincII
      >< HincII
      >< HphI
      >< EcoRI
      >< MseI
      >< Tru9I

```

FIGURA 13. 42

```

TGTGTGACACT GAAAATAACA CAGAATTCAC CAGAGTTAAT GCAAAACCTC CACCAGGTGA CCAGTTTAAA
18350      18360      18370      18380      18390      18400      18410

                >< ScrFI
                >< MvaI
                >< EcoRII
                >< Ecl136I
                >< DsaV
                >< BstOI
                >< BstNI
                >< BsiLI
                >< BsaJI
                >< NlaIII
                >< ApyI
                >< RsaI
                DdeI ><
                > < Tru9I>< Csp6I
                > < MseI >< AfaI
CATCTTATAC CACTCATGTA TAAAGGCTTG CCCTGGAATG TAGTGCGTAT TAAGATAGTA CAAATGCTCA
18420      18430      18440      18450      18460      18470      18480

                >< NlaIII
                >< HinPII
                >< Tth111I
                >< Hin6I
                >< HinfI
                > < HhaI
                >< AspI
                >< P1eI
                > < CfoI
                >< AluI
GTGATACACT GAAAGGATTG TCAGACAGAG TCGTGTTTCTT CCTTTGGGCG CATGGCTTTG AGCTTACATC
18490      18500      18510      18520      18530      18540      18550

                >< SinI
                >< Sau96I
                >< NspIV
                >< NspHII
                >< Eco47I
                >< Cfr13I
                >< ScaI
                >< BsiZI
                >< RsaI
                >< Bme18I
                >< Csp6I
                >< AvaII
                >< MaeII
                >< AfaI
                >< AsuI
                >< AflIII
                >< MaeIII>< MaeII
AATGAAGTAC TTTGTCAAGA TTGGACCTGA AAGAACGTGT TGTCTGTGTG ACAAACGTGC AACTTGCTTT
18560      18570      18580      18590      18600      18610      18620

                > < TfiI
                >< Tth111I
                > < HinfI
                > < AspI
TCTACTTCAT CAGATACTTA TGCCTGCTGG AATCATTCTG TGGGTTTTGA CTATGTCTAT AACCCATTTA
18630      18640      18650      18660      18670      18680      18690

                >< ScrFI
                RsaI ><
                >< MvaI
                >< EcoRII
                Ecl136I ><
                >< DsaV
                Csp6I ><
                BstXI ><
                > < MaeIII
                > < EcoO65I
                > < Eco91I
                > < BstPI
                >< Eco57I>
                >< BstEII
                >< MaeIII
                >< NlaIII
                >< AfaI ><
TGATTGATGT TCAGCAGTGG GGCTTTACGG GTAACCTTCA GAGTAACCAT GACCAACATT GCCAGGTACA
18700      18710      18720      18730      18740      18750      18760

                >< SfaNI
                >< RmaI
                >< NspI
                >< NspHI

```

FIGURA 13.43

```

                >< NlaIII
                >< MaeI
    >< NlaIII    >< BspWI
    > < AflIII
TGGAAATGCA CATGTGGCTA GTTGTGATGC TATCATGACT AGATGTTTAG CAGTCCATGA GTGCTTTGTT
    18770      18780      18790      18800      18810      18820      18830

    >< ThaI
    >< MvnI
    >< HinPII
    >< Hin6I
    >< HhaI
    >< CfoI
    >< BstUI
    >< Bsp50I
    >< AccII
                >< EcoNI> < MnlI
                >< BslI
                >< BsiYI
                >< Tru9I
                >< DdeI >< MseI
AAGCGCGTTG ATTGGTCTGT TGAATACCCT ATTATAGGAG ATGAACTGAG GGTAAATTCT GCTTGCAGAA
    18840      18850      18860      18870      18880      18890      18900

    >< RsaI
    >< Csp6I
    >< AfaI    >< NlaIII
                >< BspWI
                >< MboII
                > < NlaIII
                >< BsrI >< BspHI
AAGTACAACA CATGGTTGTG AAGTCTGCAT TGCTTGCTGA TAAGTTTCCA GTTCTTCATG ACATTGGAAA
    18910      18920      18930      18940      18950      18960      18970

                >< SauI
                >< MstII
                >< Eco8II
                >< DdeI
                >< CvnI
                >< Bsu36I
                >< Bse2II
                >< AxyI
                >< AocI
                >< MnlI
                >< SfaNI
                >< Bpu1102I
                NlaIII ><
                >< EspI
                >< Eco57I MaeIII ><
                >< DdeI
                >< CelII
                >< Bpu1102I
TCCAAAGGCT ATCAAGTGTG TGCCTCAGGC TGAAGTAGAA TGGAAGTTCT ACGATGCTCA GCCATGTAGT
    18980      18990      19000      19010      19020      19030      19040

                >< MnlI
                >< Ksp632I
    >< HindIII
    >< AluI    >< MboII
    >< Eam1104I
GACAAAGCTT ACAAATAGA GGAAGTCTTC TATTCTTATG CTACACATCA CGATAAATTC ACTGATGGTG
    19050      19060      19070      19080      19090      19100      19110

                >< Sau3AI
                >< NdeII
                >< MboI
    >< MaeII> < MaeIII
    >< DpnII
    >< DpnI
    >< BspAI
                >< MaeIII >< Bsp143I
                >< MunI
                >< HinfI >
                >< OrdI ><
TTTGTGTTGTT TTGGAATTGT AACGTTGATC GTTACCCAGC CAATGCAATT GTGTGTAGGT TTGACACAAG
    19120      19130      19140      19150      19160      19170      19180

                >< Zsp2I ><
                >< SphI
                > < Ppu10I
                >< PaeI
                >< NspI
                >< NspHI
                >< NlaIII
                >< Mph1103I ><
    >< ScrFI
    >< MvaI
    >< EcoRII

```

FIGURA 13. 44

FIGURA 13. 45

```

>< XhoII
>< Sau3AI
>< NdeII
>< MflI
>< MboI
>< DpnII
  >< DpnI
>< BstYI
>< BspAI
  >< BspI43I
  >< BglII
GAGATCTTTG AAAATAAGAC AACACTTCCT GTTAATGTTG CATTTGAGCT TTGGGCTAAG CGTAACATTA
19680      19690      19700      19710      19720      19730      19740

                                >< MaeIII
                                >< EspI
                                >< DdeI Tru9I ><
                                >< CelIIMseI ><
                                >< BpuII02I
                                >< AluI
                                >< Fnu4HI
                                >< EcoRV
                                >< Eco32I
  >< BsrI      >< Tru9I      >< BbvI      >< MseI
AACCAGTGCC AGAGATTAAG ATACTCAATA ATTTGGGTGT TGATATCGCT GCTAATACTG TAATCTGGGA
19750      19760      19770      19780      19790      19800      19810

                                >< NspI
                                >< NspHI
                                >< NlaIII
                                >< BsgI
                                >< AflIII
CTACAAAAGA GAAGCCCCAG CACATGTATC TACAATAGGT GTCTGCACAA TGA CTGACAT TGCCAAGAAA
19820      19830      19840      19850      19860      19870      19880

  >< DdeI>< MboII
CCTACTGAGA GTGCTTGTTT TTCACCTACT GTCTTGTTTG ATGGTAGAGT GGAAGGACAG GTAGACCTTT
19890      19900      19910      19920      19930      19940      19950

                                SinI ><
                                Sau96I ><
                                NspIV ><
                                NspHII ><
                                NlaIV ><
                                Eco47I ><
                                CfrI3I ><
                                >< BslI
                                Bsi2I ><
                                >< BsiYI
                                BscBI ><
                                Bmel8I ><
                                AvaII ><
                                AsuI ><
                                >< Tru9I
                                >< MseI
TTAGAAACGC CCGTAATGGT GTTTTAATAA CAGAAGGTTT AGTCAAAGGT CTAACACCTT CAAAGGGACC
19960      19970      19980      19990      20000      20010      20020

                                >< VspI
                                >< Tru9I
                                >< PflI
                                >< MseI
  >< RmaI
  >< NheI      >< MaeIII
  >< MaeI      >< AsnI      >< TfiI
>< HgaI>< AluI >< HinfI>< AseI      >< HinfI      >< MseI
AGCACAAGCT AGCGTCAATG GAGTCACATT AATTGGAGAA TCAGTAAAAA CACAGTTTAA CTACTTTAAG
20030      20040      20050      20060      20070      20080      20090

                                >< DdeI      >< MnlI      Tru9I ><
                                >< BsmAI      >< DdeI

```

FIGURA 13. 46

```

>< AccI                               >< Alw26I >< BfrIMseI ><
AAAGTAGACG GCATTATTCA ACAGTTGCCT GAAACCTACT TTACTCAGAG CAGAGACTTA GAGGATTTTA
  20100      20110      20120      20130      20140      20150      20160

                                >< TthHB8I
                                >< TaqI
                                    >< SstI
                                    >< SduI
                                    >< SacI
                                > < PaeR7I
                                > < NspIII
                                    >< NspII
                                    >< HgiAI
                                > < Eco88I
                                > < XhoI>< Eco24I
                                    >< Ecl136II
                                > < SlaI>< Bsp1286I
                                > < CcrI>< BmyI
                                > < BcoI>< BanII
                                > < Ama87I
                                > < AvaI>< Alw21I
                                >< AluI
                                >< EcoRI >< FokIAluI ><
XhoI ><
TthHB8I >
TaqI >
SlaI ><
PaeR7I ><
NspIII ><
MnlI
Eco88I ><
CcrI ><
BspWI ><
BcoI ><
> < BcgI/a
AvaI ><
Ama87I ><
>< XcmI
>< Sau3AI
>< NdeII
>< MboI
>< DpnII
>< DpnI
>< BspAI
>< Bsp143I
AGCCCAGATC ACAAATGGAA ACTGACTTTC TCGAGCTCGC TATGGATGAA TTCATACAGC GATATAAGCT
  20170      20180      20190      20200      20210      20220      20230

                                >< TthHB8I
                                >< TaqI
                                >< SfuI
                                >< NspV
                                >< LspI
                                >< Csp45I
                                >< BstBI
                                >< Bsp119I
                                >< BsiCI
                                >< Bpu14I
                                >< AsuII >< BcgI
                                >< MboII
                                >< BbsI
                                >< NlaIII >< AciIMseI ><
CGAGGGCTAT GCCTTCGAAC ACATCGTTTA TGGAGATTTC AGTCATGGAC AACTTGGCGG TCTTCATTTA
  20240      20250      20260      20270      20280      20290      20300

                                >< HphI
                                >< HinP1I
                                >< Hin6I
                                >< EspI > < HhaI >< TfiI
                                >< DdeI >< HaeII
                                >< CelII >< Eco47III >< Tru9I
                                >< Bpu1102I > < CfoI >< HinfI >< MseI
                                >< BfrI >< Bsp143II >< MnlI
ATGATAGGCT TAGCCAAGCG CTCACAAGAT TCACCACTTA AATTAGAGGA TTTTATCCCT ATGGACAGCA
  20310      20320      20330      20340      20350      20360      20370

                                >< MstI
                                >< HinP1I
                                >< Hin6I
                                >< HhaI
                                >< FspI
                                >< FdiII
                                >< CfoI
                                >< SfaNI >< AviII
                                >< MboII
                                >< DpnII ><
                                >< DpnI ><
                                >< BspAI ><
                                >< Bsp143I ><
                                >< Sau3AI ><
                                >< NdeII ><
                                >< MboI ><
                                >< DpnII ><
                                >< DpnI ><
                                >< BspAI ><
                                >< Bsp143I ><
CAGTGAAAAA TTACTTCATA ACAGATGCGC AAACAGGTTT ATCAAAATGT GTGTGTCTTG TGATTGATCT
  20380      20390      20400      20410      20420      20430      20440

                                >< TthHB8I

```

FIGURA 13. 47


```

>< Tth111I
>< TaqI
>< AspI > < MaeIII MaeIII ><
TTTACTTGAT GACTTTGTCTG AGATAATAAA GTCACAAGAT TTGTCAGTGA TTTCAAAAGT GGTCAGGTT
20450 20460 20470 20480 20490 20500 20510

>< NspI
>< NspHI
>< NlaIII
>< FokI
>< MunI > < NlaIII >< AflIII
ACAATTGACT ATGCTGAAAT TTCATTATG CTTTGGTGTA AGGATGGACA TGTTGAAACC TTCTACCCAA
20520 20530 20540 20550 20560 20570 20580

>< SfaNI
>< ScrFI
>< MvaI
>< EcoRII
>< Ecl136I
>< DsaV
>< BstOI >< SfaNI
>< BstNI >< RsaI BspWI ><
>< BsiLI > < Csp6I BsmI >
>< BspWI >< ApyI >< AfaI BscCI ><
AACTACAAGC AAGTCAAGCG TGGCAACCAG GTGTTGCGAT GCCTAACTTG TACAAGATGC AAAGAATGCT
20590 20600 20610 20620 20630 20640 20650

>< Eco57I >< MaeIII >< HphI
TCTTGAAAAG TGTGACCTTC AGAATTATGG TGAAAATGCT GTTATACCAA AAGGAATAAT GATGAATGTC
20660 20670 20680 20690 20700 20710 20720

> < RsaI
>< Csp6I
>< Bst1107I >< Tru9I >< AluI
>< AccI >< MseI > < AfaINlaIII ><
GCAAAGTATA CTCAACTGTG TCAATACTTA AATACACTTA CTTTAGCTGT ACCCTACAAC ATGAGAGTTA
20730 20740 20750 20760 20770 20780 20790

>< ScrFI
>< RsaI
>< MvaI
>< EcoRII >< Nsp8II
>< Ecl136I >< SduI
> < Csp6I >< NspII
>< BstOI >< PvuII>< HgiAI
>< BstNI >< DdeI
>< BsiLI >< Psp5I>< Bsp1286I
>< ApyI >< AluI >< BmyI
>< DsaV>< AfaI >< Alw21I
TTCACTTTGG TGCTGGCTCT GATAAAGGAG TTGCACCAGG TACAGCTGTG CTCAGACAAT GGTGCGCAAC
20800 20810 20820 20830 20840 20850 20860

>< XhoII
>< Tru9I
>< Sau3AI
>< NdeII
>< TthHB8I >< MseI
>< MflI
>< MboI
>< MamI
>< DpnII
>< TfiI >< DpnI

```

FIGURA 13. 48

```

                >< BstYI                > < TfiI
                >< BspAI                > < HinfI
                >< HinfI>< Bsp143I        >< Esp3I        >< Tru9I
                >< BsiBI        >< Tth111I    >< BsmBI        >< MseI
                >< BsaBI        >< BsmAI        > < BsmAI
                >< BsrI        >< TaqI >< BglII    >< AspI        >< Alw26I >< HgaI> < Alw26I
TGGCACACTA CTTGTCGATT CAGATCTTAA TGACTTCGTC TCCGACGCAG ATTCTACTTT AATTGGAGAC
    20870      20880      20890      20900      20910      20920      20930

                >< StyI
                >< SinI
                >< Sau96I
                > < SinI                >< RmaI
                > < Sau96I            >< NspIV
                >< PssI                NspHII ><
                >< Psp5II            >< MaeI
                > < PpuMI            >< EcoT14I
                > < NspIV            >< Eco47I
                >< NspHII            >< Eco130I
                >< NlaIV            >< Cfr13I
                > < EcoO109I        >< BssT1I
                > < Eco47I            >< BsiZI
                > < DraII            >< BsaJI
                > < Cfr13I            >< Bme18I
                > < BsiZI            >< BlnI
                >< BscBI            >< AvrII
                >< RsaI                >< Bme18I        >< AvaII
                > < Csp6I            >< AvaII            >< AsuI
                >< AfaI                >< AsuI            AflIII ><
TGTGCAACAG TACATACGGC TAATAAATGG GACCTTATTA TTAGCGATAT GTATGACCCT AGGACCAAAC
    20940      20950      20960      20970      20980      20990      21000

                >< NspI
                >< NspHI
                >< NlaIII >< P1eI                RmaI ><
                >< MaeIII        >< HinfI            MaeI ><
ATGTGACAAA AGAGAATGAC TCTAAAGAAG GGTTTTTTAC TTATCTGTGT GGATTTATAA AGCAAAAAC
    21010      21020      21030      21040      21050      21060      21070

                >< ScrFI
                >< MvaI
                >< EcoRII
                >< Ecl136I
                >< DsaV
                >< BstOI                Sau96I >
                >< BstNI                NspIV >
                >< BsiLI                Cfr13I >
                >< BsaJI                BsiZI >
                >< BsaJI        >< SfcI                >< BsmI        >< BsmI        >< AsuI >
                >< ApyI                > < AluI        >< BscCI        >< BscCIHindIII ><>< AluI
AGCCCTGGGT GGTTCATAG CTGTAAAGAT AACAGAGCAT TCTTGGAATG CTGACCTTTA CAAGCTTATG
    21080      21090      21100      21110      21120      21130      21140

                >< Zsp2I
                >< Ppu10I
                >< NsiI
                >< Mph1103I    Tru9I ><
                >< BsuRI                >< MaeIII        >< EcoT22I        >< MseI
                >< BshI        >< NlaIII>< AluI    >< BcgI        >< AvaIII >< SfaNIBcgI/a ><
GGCCATTTCT CATGGTGGAC AGCTTTTGTT ACAAATGTAA ATGCATCATC ATCGGAAGCA TTTTAAATTG
    21150      21160      21170      21180      21190      21200      21210

```

FIGURA 13. 49

```

                                >< Zsp2I
                                >< SphI
                                >< Ppu10I
                                >< PaeI
                                >< NspI
                                >< NspHI
                                >< NsiI
                                >< NlaIII
                                > < NlaIII
                                >< Mph1103I
                                >< EcoT22I
                                > < AvaIII >< MnlI
GGGCTAACTA TCTTGGCAAG CCGAAGGAAC AAATTGATGG CTATACCATG CATGCTAACT ACATTTTCTG
21220      21230      21240      21250      21260      21270      21280

                                Tru9I ><
                                >< Tru9I
                                >< GsuI      MseI ><
                                >< BsrI      >< MseI
                                >< BpmI      MnlI ><
                                >< BbsI      >< NlaIII >< MnlI
GAGGAACACA AATCCTATCC AGTTGTCTTC CTATTCACCTC TTTGACATGA GCAAATTTCC TCTTAAATTA
21290      21300      21310      21320      21330      21340      21350

                                >< Tru9I
                                >< MseI
                                >< Esp4I> < TfiI
                                >< BsmAI      Ksp632I ><
                                >< Alw26I      >< MboII >< EarI
                                >< AflII> < HinfI      Eam1104I ><
AGAGGAACTG CTGTAATGTC TCTTAAGGAG AATCAAATCA ATGATATGAT TTATTCTCTT CTGGAAAAAG
21360      21370      21380      21390      21400      21410      21420

                                >< Tru9I
                                >< MseI
                                >< HindII
                                >< HincII
                                >< HpaI AflIII >
GTAGGCTTAT CATTAGAGAA AACAACAGAG TTGTGGTTTC AAGTGATATT CTTGTTAACA ACTAAACGAA
21430      21440      21450      21460      21470      21480      21490

                                >< VneI
                                >< SnoI
                                >< SduI
                                >< NspII
                                >< HpaII
                                >< HgiAI
                                >< HapII
                                >< Cfr10I
                                >< Bsp1286I
                                >< MspI>< BmyI
                                >< ApaLI
                                >< Alw44I
                                >< MaeI >< MaeIII >< AgeI >< Alw21I
                                >< NspI
                                >< NspHI
                                >< NlaIII
CATGTTTATT TTCTTATTAT TTCTTACTCT CACTAGTGGT AGTGACCTTG ACCGGTGCAC CACTTTTGAT
21500      21510      21520      21530      21540      21550      21560

                                > < AluI
                                >< MnlI
GATGTTCAAG CTCCTAATTA CACTCAACAT ACTTCATCTA TGAGGGGGGT TTACTATCCT GATGAAATTT
21570      21580      21590      21600      21610      21620      21630

                                >< Sau3AI

```

FIGURA 13. 50

```

>< NdeII
>< MboI
>< DpnII
  >< DpnI          >< Tru9I
>< BspAI          >< MseI > < MboII
  >< Bsp143I       >< DdeI          >< MaeIII
TTAGATCAGA CACTCTTTAT TTAAGTCAGG ATTTATTTCT TCCATTTTAT TCTAATGTTA CAGGGTTTCA
    21640      21650      21660      21670      21680      21690      21700

  >< VspI
  >< Tru9I
  >< MseI
  >< AsnI          >< Tru9I          >< FokI
  >< AseI >< MaeII >< MseI >< BbvI          > < Fnu4HI
TACTATTAAT CATACGTTTG GCAACCCTGT CATACCTTTT AAGGATGGTA TTTATTTTGC TGCCACAGAG
    21710      21720      21730      21740      21750      21760      21770

          >< BslI
          >< DsaI>< BsiYI          >< NlaIII
          >< BsaJI          > < MaeIII
AAATCAAATG TTGTCCGTGG TTGGGTTTTT GGTCTACCA TGAACAACAA GTCACAGTCG GTGATTATTA
    21780      21790      21800      21810      21820      21830      21840

          >< NspI
>< Tru9I          >< NspHI
>< MseI          >< NlaIII
>< HphI          >< MaeIII          >< MaeIII
TTAACAATTC TACTAATGTT GTTATACGAG CATGTAACCT TGAATTGTGT GACAACCCTT TCTTTGCTGT
    21850      21860      21870      21880      21890      21900      21910

  >< StyI          >< Zsp2I
    >< NlaIII          >< Tru9I
  >< NcoI >< RsaI          >< Ppu10I TthHB8I ><
  >< EcoT14I          >< NsiI          >< TaqI
  >< Eco130I          >< MseI          SfaNI ><
  >< DsaI>< Csp6I          >< Mph1103I RsaI ><
  >< BssT1I          >< TthHB8I >< EcoT22I Csp6I ><
  >< BsaJI>< AfaI          >< TaqI >< AvaIII AfaI ><
TTCTAAACCC ATGGGTACAC AGACACATAC TATGATATTC GATAATGCAT TTAATTGCAC TTTTCGAGTAC
    21920      21930      21940      21950      21960      21970      21980

          >< Tru9I
          >< MseI
          >< DraI
ATATCTGATG CCTTTTCGCT TGATGTTTCA GAAAAGTCAG GTAATTTTAA ACACTTACGA GAGTTTGTGT
    21990      22000      22010      22020      22030      22040      22050

          >< Sau3AI
          >< NdeII
          >< MboI
          >< DpnII
>< Tru9I          >< DpnI
>< MseI          >< BspAI
  >< DraI          >< SfcI Bsp143I ><
TTAAAAATAA AGATGGGTTT CTCTATGTTT ATAAGGGCTA TCAACCTATA GATGTAGTTC GTGATCTACC
    22060      22070      22080      22090      22100      22110      22120

          >< Tru9I
  >< Tru9I          > < Tru9I          >< MseI
  >< MseI          > < MseI          >< MnlI
TTCTGGTTTT AACACTTTGA AACCTATTTT TAAGTTGCCT CTTGGTATTA ACATTACAAA TTTTAGAGCC
    22130      22140      22150      22160      22170      22180      22190

```

FIGURA 13. 51

```

> < SduI>< SfcI
>< PvuII
>< Psp5I
> < NspII
>< NspBII
> < MaeII > < Fnu4HI
> < Bsp1286I >< PstI Tru9I >
>< BmyI>< Fnu4HI MseI >
>< HphI >< BspMI >< BbvI >< AluI >< BbvI
ATTCTTACAG CCTTTTCACC TGCTCAAGAC ATTTGGGGCA CGTCAGCTGC AGCCTATTTT GTTGGCTATT
22200 22210 22220 22230 22240 22250 22260

>< SfaNI
>< RsaI
> < Csp6I
>< AfaI >< AlwNI
>< DraI
TAAAGCCAAC TACATTTATG CTCAAGTATG ATGAAAATGG TACAATCACA GATGCTGTTG ATTGTTCTCA
22270 22280 22290 22300 22310 22320 22330

> < Tru9I
> < MseI
>< AluI
AAATCCACTT GCTGAAGTCA AATGCTCTGT TAAGAGCTTT GAGATTGACA AAGGAATTTA CCAGACCTCT
22340 22350 22360 22370 22380 22390 22400

>< SauI
>< MstII
>< Eco81I
>< DdeI
>< CvnI
>< Bsu36I
>< Bse21I
>< AxyI >< TfiI
>< MnlI >< AocI >< MnlI >< HinfI >< SspI >< MnlI
AATTTCAAGG TTGTTCCCTC AGGAGATGTT GTGAGATTCC CTAATATTAC AAAGTTGTGT CCTTTTGAG
22410 22420 22430 22440 22450 22460 22470

>< Zsp2I
>< Ppu10I
>< NsiI
> < NlaIII
>< Mph1103I
>< EcoT22I
>< Tru9I
>< MseI >< AvaIII
AGGTTTTTAA TGCTACTAAA TTCCCTTCTG TCTATGCATG GGAGAGAAAA AAAATTTCTA ATTGTGTTGC
22480 22490 22500 22510 22520 22530 22540

>< SduI
>< NspII
>< HgiAI
>< Bsp1286I
>< BmyI >< Tru9I
>< Alw21I >< MseI DdeI ><
TGATTACTCT GTGCTCTACA ACTCAACATT TTTTCAACC TTAAAGTGCT ATGGCGTTTC TGCCACTAAG
22550 22560 22570 22580 22590 22600 22610

>< Sau3AI
>< NdeII
>< MboI
>< DpnII
>< DpnI

```

FIGURA 13. 52

ES 2 396 127 T3

```

    >< BspAI
    >< Bsp143I
    TTGAATGATC TTTGCTTCTC CAATGTCTAT GCAGATTCTT TTGTAGTCAA GGGAGATGAT GTAAGACAAA
    22620      22630      22640      22650      22660      22670      22680

    >< ScrFI
    >< MvaI
    >< HinPII
    >< Hin6I
    >< HhaI
    >< HaeII
    >< EcoRII
    >< Ecl136I
    >< DsaV
    >< CfoI
    >< BstOI
    >< BstNI
    >< Bsp143II
    >< BsiLI
    >< ApyI      > < BsrI
    TAGCGCCAGG ACAAACGGT GTTATTGCTG ATTATAATTA TAAATTGCCA GATGATTTCa TGGGTTGTGT
    22690      22700      22710      22720      22730      22740      22750

    >< SfaNI
    >< RmaI
    >< MaeI
    CCTTGCTTGG AATACTAGGA ACATTGATGC TACTTCAACT GGTAATTATA ATTATAAATA TAGGTATCTT
    22760      22770      22780      22790      22800      22810      22820

    >< Sau96I
    >< Pali
    >< NspIV
    > < HindIII
    >< HaeIII
    >< EcoO109I
    >< DraII
    >< DdeI
    >< Cfr13I
    >< BsuRI
    >< BsiZI
    >< BshI
    >< BfrI >< PssI
    >< NlaIII >< AsuI>< BsmAI
    >< AluI      >< Alw26I
    AGACATGGCA AGCTTAGGCC CTTTGAGAGA GACATATCTA ATGTGCCTTT CTCCCCTGAT GGCAAACCTT
    22830      22840      22850      22860      22870      22880      22890

    >< Tru9I
    >< Pali
    >< MscI
    >< HaeIII
    >< EaeI>< MseI
    >< Tru9I      >< BsuRI
    >< MseI      >< BshI
    >< BspMI      >< BalI
    GCACCCCAACC TGCTCTTAAT TGTTATTGGC CATTAAATGA TTATGGTTTT TACACCACTA CTGGCATTGG
    22900      22910      22920      22930      22940      22950      22960

    Sau96I ><
    >< PaliNspIV ><
    > < MspI NspHII ><
    >< HaeIII

```

FIGURA 13. 53

```

                                > < HpaII Eco47I ><
                                >< DsaI
                                > < HapII Cfr13I ><
                                >< BsuRISinI ><
                                >< GdiIII BsiZI ><
                                >< BsaJI
                                >< ScaI
                                >< RsaI
                                >< Csp6I
                                >< AfaI
                                >< Tru9I
                                >< MseI >< Cfr10I
                                >< DraI
                                >< EaeI Bmel8I ><
                                >< AvaII ><
                                >< BshI AsuI ><

CTACCAACCT TACAGAGTTG TAGTACTTTC TTTTGAAC TTAAATGCAC CGGCCACGGT TTGTGGACCA
  22970      22980      22990      23000      23010      23020      23030

                                >< Tru9I
                                >< Tru9I
                                >< PleI
                                >< MseI
                                >< RsaI
                                >< Csp6I
                                >< BsrI ><
                                >< BsrI
                                >< MseI
                                >< HinfI
                                >< AfaI
                                >< MseI>< BsrI
                                >< MseI
                                >< HinfI
                                >< AfaI

AAATTATCCA CTGACCTTAT TAAGAACCAG TGTGTCAATT TTAATTTTAA TGGACTCACT GGTACTGGTG
  23040      23050      23060      23070      23080      23090      23100

                                >< Tru9I
                                >< MseI
                                >< MboII
                                >< HpaI
                                >< HindII
                                >< HincII
                                >< Pali
                                >< HaeIII
                                >< GdiIII
                                >< EaeI
                                >< BsuRI
                                >< BshI
                                TfilI ><
                                HinfI ><

TGT TAACTCC TTCTTCAAAG AGATTTC AAC CATTTC AACA ATTTGGCCGT GATGTTTCTG ATTTCACTGA
  23110      23120      23130      23140      23150      23160      23170

                                > < XhoII
                                >< TthHB8I
                                >< TaqI
                                > < Sau3AI
                                > < NdeII
                                > < MflI
                                > < MboI
                                > < DpnII
                                >< DpnI
                                > < BstYI
                                > < BspAI
                                > < SspI
                                >< AlwI >< Bsp143I
                                >< HphI

TTCCGTTTCGA GATCCTAAAA CATCTGAAAT ATTAGACATT TCACCTTGCT CTTTGGGGG TGTAAGTGTA
  23180      23190      23200      23210      23220      23230      23240

                                >< ScrFI
                                >< MvaI
                                >< EcoRII
                                >< Ecl136I
                                >< DsaV
                                >< BstOI
                                >< BstNI
                                >< BsiLI
                                >< ApyI
                                >< Tru9I
                                >< MseI
                                >< HpaI
                                >< HindII
                                >< Eco57I
                                >< BsgI
                                >< HincII

ATTACACCTG GAACAAATGC TTCA1CTGAA GTTGCTGTTC TATATCAAGA TGT TAACTGC ACTGATGTTT
  23250      23260      23270      23280      23290      23300      23310

                                >< Sau3AI
                                >< NlaIII
                                >< NdeII
                                >< MboI
                                >< DpnII
                                >< DpnI
                                >< HinfII

```

FIGURA 13. 54

```

>< BspWI                >< Hin6I
    >< BspAI                > < HhaI                PleI ><
>< SfcI                >< Bsp143I        >< AluI> < CfoI                >< BsrI
CTACAGCAAT TCATGCAGAT CAACTCACAC CAGCTTGGCG CATATATTCT ACTGGAAACA ATGTATTCCA
    23320        23330        23340        23350        23360        23370        23380

                                >< TthHB8I
                                >< TaqI
                                >< SalI
                                >< RtrI
                                >< NspI
                                >< EspI >< NspHI
                                >< DdeI >< NlaIII
                                >< CelII >< HindII
                                >< Bpu1102I>< HincII
>< HinfI                >< AluI        >< AccI
GACTCAAGCA GGCTGTCTTA TAGGAGCTGA GCATGTCGAC ACTTCTTATG AGTGCGACAT TCCTATTGGA
    23390        23400        23410        23420        23430        23440        23450

                                > < SnaBI
                                >< ScaI
                                >< RsaI
                                >< RmaI
                                >< MaeII >< MaeI
                                > < Eco105I
                                >< Csp6I
                                > < BsaAI
                                >< AfaI
>< AluI                >< MaeI
GCTGGCATT TGTGCTAGTTA CCATACAGTT TCTTTATTAC GTAGTACTAG CCAAAAATCT ATTGTGGCTT
    23460        23470        23480        23490        23500        23510        23520

                                >< MunI
ATACTATGTC TTTAGGTGCT GATAGTTCAA TTGCTTACTC TAATAACACC ATTGCTATAC CTACTAACTT
    23530        23540        23550        23560        23570        23580        23590

                                RsaI ><
                                >< MnlI
                                Csp6I ><
                                AfaI ><
>< SfcI
TTCAATTAGC ATTACTACAG AAGTAATGCC TGTTTCTATG GCTAAACCT CCGTAGATTG TAATATGTAC
    23600        23610        23620        23630        23640        23650        23660

    > < TfiI
    > < HinfI
>< AciI                > < AluI
ATCTGCGGAG ATTCTACTGA ATGTGCTAAT TTGCTTCTCC AATATGGTAG CTTTGCACA CAACTAAATC
    23670        23680        23690        23700        23710        23720        23730

>< VneI
    >< SduI
    >< NspII
    >< HgiAI                >< PmlI
>< SnaI>< DdeI                >< Sau3AI                >< PmaCI
    >< Bsp1286I                >< NdeII                >< MaeII
    >< BmyI                >< MboI                >< Eco72I
    >< BbvI                >< DpnI                >< BsaAI
>< ApaLI                >> Bsp143I >< BbrPI
>< Alw44I                >< DpnII >< AlwI
    >< Alw21I                >< Fnu4HI >< BspAI >< AflIII
GTGCACTCTC AGGTATTGCT GCTGAACAGG ATCGCAACAC ACGTGAAGTG TTCGCTCAAG TCAAACAAAT
    23740        23750        23760        23770        23780        23790        23800

```

FIGURA 13. 55

ES 2 396 127 T3

```

>< RsaI
>< Csp6I
>< AfaI
GTACAAAACC CCAACTTTGA AATATTTTGG TGGTTTAAAT TTTTCACAAA TATTACCTGA CCCTCTAAAG
23810      23820      23830      23840      23850      23860      23870

>< MnlI
>< MnlI
>< DdeI >< MnlI
>< Tru9I >< SfaNI >< HphI NlaIII ><
>< MseI >< MaeIII BspHI ><
CCAACTAAGA GGTCTTTTAT TGAGGACTTG CTCTTTAATA AGGTGACACT CGCTGATGCT GGCTTCATGA
23880      23890      23900      23910      23920      23930      23940

>< XhoII
>< Sau3AI
>< StyI >< RmaI
>< RmaI >< NdeII
>< MaeI >< MflI
>< EcoT14I >< MboI >< MstI
>< Eco130I >< MaeI >< HinPII
>< BssTII >< VspI >< DpnII >< Hin6I
>< BsmI >< HphI> < DpnI >< HhaI
>< BscCI >< Tru9I >< BstYI >< FspI
>< BsaJI >< MseI >< BspAI >< FdiII
>< BlnI >< AsnI >< Bsp143I >< CfoI
>< AvrII >< AseI >< BglII >< AviII
AGCAATATGG CGAATGCCTA GGTGATATTA ATGCTAGAGA TCTCATTTGT GCGCAGAAGT TCAATGGACT
23950      23960      23970      23980      23990      24000      24010

>< RmaIRsaI ><
>< MnlI >< Fnu4HI >< Fnu4HI Csp6I ><
>< BspWI >< BbvI >< BbvI >< BspWI >< MaeIAfaI ><
TACAGTGTTG CCACCTCTGC TCACTGATGA TATGATTGCT GCCTACACTG CTGCTCTAGT TAGTGGTACT
24020      24030      24040      24050      24060      24070      24080

>< MboII
>< HinPII
>< Hin6I
>< HhaI
>< HaeII
>< Fnu4HI >< Ksp632I
>< CfoI >< EarI
>< FokI >< BspWI >< Eam1104I
>< BbvI >< Bsp143II
GCCACTGCTG GATGGACATT TGGTGCTGGC GCTGCTCTTC AAATACCTTT TGCTATGCAA ATGGCATATA
24090      24100      24110      24120      24130      24140      24150

>< MaeIII
>< Tru9I ><
>< MseI ><
GGTTCAATGG CATTGGAGTT ACCCAAAATG TTCTCTATGA GAACCAAAAA CAAATCGCCA ACCAATTTAA
24160      24170      24180      24190      24200      24210      24220

>< MaeII ><
>< TfiI >< Fnu4HI
>< HinfI >< BbvI >< AluI
CAAGGCGATT AGTCAAATTC AAGAATCACT TACAACAACA TCAACTGCAT TGGGCAAGCT GCAAGACGTT
24230      24240      24250      24260      24270      24280      24290

>< Tru9I
>< MseI
>< HpaI >< DdeI
>< HindII >< BsmI >< Tru9I >< Tru9I >< BfrI
>< HincII>< BscCI >< MseI >< MseI >< AluI

```

FIGURA 13. 56

```

GTTAACCAGA ATGCTCAAGC ATTAAACACA CTTGTTAAAC AACTTAGCTC TAATTTTGGT GCAATTTCAA
24300      24310      24320      24330      24340      24350      24360

      >< ThaI
      >< SpoI
      >< NruI
      >< MvnI
      >< BstUI      >< TthHB8I
      >< Bsp68I      >< TaqI      >< RsaI
      >< EcoRV      >< Bsp50I      >< MnlI      >< Csp6I      >< Tru9I
      >< Eco32I      >< AccII      >< MnlI      >< AciI      >< AfaI      >< MseI
GTGTGCTAAA TGATATCCTT TCGCGACTTG ATAAAGTCGA GGCGGAGGTA CAAATTGACA GGTTAATTAC
24370      24380      24390      24400      24410      24420      24430

      >< MaeIII      >< BbvI      >< Fnu4HI      BbvI      ><
AGGCAGACTT CAAAGCCTTC AAACCTATGT AACACAACAA CTAATCAGGG CTGCTGAAAT CAGGGCTTCT
24440      24450      24460      24470      24480      24490      24500

      >< Fnu4HI      >< HindII
      >< BspWI      >< DdeI      >< HincII
GCTAATCTTG CTGCTACTAA AATGTCTGAG TGTGTTCTTG GACAATCAAA AAGAGTTGAC TTTGTGGAA
24510      24520      24530      24540      24550      24560      24570

      > < NspI
      > < NspHI
      > < NlaIII
      >< MaeIII
      >< NlaIII      >< MaeII
      >< MboII      >< FokI
      >< Fnu4HI      >< BbsI      BsaAI      ><
      >< AciI      >< BbvI      >< AflIII
AGGGCTACCA CCTTATGTCC TTCCACAAG CAGCCCGCA TGGTGTGTC TTCCTACATG TCACGTATGT
24580      24590      24600      24610      24620      24630      24640

      >< ScrFI
      >< MvaI
      >< EcoRII
      >< Ecl136I
      >< BstOI
      >< BstNI      >< HinfI
      >< MnlI      >< BslI      >< HinfI
      >< DsaV      >< BsiYI      >< HhaI
      >< BsiLI      >< HaeII
      >< BsaJI      >< HphI      >< CfoI      >< NlaIII
      >< ApyI      >< Bsp143II      >< BspHI      EcoNI      ><
GCCATCCAG GAGAGGAACT TCACCACAGC GCCAGCAATT TGTCATGAAG GCAAAGCATA CTTCCCTCGT
24650      24660      24670      24680      24690      24700      24710

      >< MnlI
      >< BslI      >< Tru9I
      >< BsiYI      >< MseI      >< MnlI
GAAGGTGTTT TTGTGTTTAA TGGCACTTCT TGGTTTATTA CACAGAGGAA CTTCTTTTCT CCACAAATAA
24720      24730      24740      24750      24760      24770      24780

      >< DdeI      >< Tru9I
      >< BsmAI      >< SfaNI
      >< SfcI      >< Alw26I      >< MseI      >< AlwI      ><
TTACTACAGA CAATACATTT GTCTCAGGAA ATTGIGATGT CGTTATTGGC ATCATTAACA ACACAGTTTA
24790      24800      24810      24820      24830      24840      24850

      >< Sau3AI
      >< NdeII

```

FIGURA 13. 57

```

>< MboI          >< PfuI          > < ScaI
>< DpnII         >< MnlI          > < Ksp632I      > < RsaI
  >< DpnI         >< DdeI   >< HinfI      >< MboII
>< BspAI         >< BspWI          > < Eam1104I      >< Csp6I
  >< Bsp143I      >< AluI          > < EarI   > < AluI   > < AfaI   > < HphI
TGATCCTCTG CAACCTGAGC TTGACTCATT CAAAGAAGAG CTGGACAAGT ACTTCAAAAA TCATACATCA
      24860      24870      24880      24890      24900      24910      24920

  >< Sau3AI
  >< NdeII
  >< MboI
>< MamI
  >< DpnII
    >< DpnI
  >< BspAI
    >< Bsp143I
  >< BsiBI          >< Tru9I          >< HindII
  >< BsaBI          >< MseI          >< HincII      AclI ><
CCAGATGTTG ATCTTGCGCA CATTTCAGGC ATTAACGCTT CTGTCGTCAA CATTCAAAAA GAAATTGACC
      24930      24940      24950      24960      24970      24980      24990

    >< Tru9I
      > < TfiI
    >< MnlI          >< SmaI
  >< EcoNI          >< MseI
    >< BslI          > < HinfI
>< MnlI>< BsiYI      >< DraI
GCCTCAATGA GGTGCTTAAA AATTTAAATG AATCACTCAT TGACCTTCAA GAATTGGGAA AATATGAGCA
      25000      25010      25020      25030      25040      25050      25060

    >< StyI
    >< PstI
    >< HaeIII
    >< EcoT14I
    >< Eco130I
    >< BsuRI
    >< BstXI
  >< Tru9I>< BshI          NlaIII ><
  >< MseI   >< BsaJI      MaeIII ><
ATATATTAAA TGGCCTTGGT ATGTTTGGCT CGGCTTCATT GCTGGACTAA TTGCCATCGT CATGGTTACA
      25070      25080      25090      25100      25110      25120      25130

    > < SphI
    > < PaeI
    >< SpeI          > < NspI
    > < RmaI        > < NspHI
    >< NlaIII       > < NlaIII
    > < MaeI        >< MnlI>< BbvI   Fnu4HI ><
ATCTTGCTTT GTTGCAATGAC TAGTTGTTGC AGTTGCCTCA AGGGTGCATG CTCTTGTTGGT TCTTGCTGCA
      25140      25150      25160      25170      25180      25190      25200

    >< FokI
    >< DdeI
>< MnlI >< PfuI>< HinfI >< BsrI
AGTTTGATGA GGATGACTCT GAGCCAGTTC TCAAGGGTGT CAAATTACAT TACACATAAA CGAAGTTATG
      25210      25220      25230      25240      25250      25260      25270

    >< Sau3AI
    >< NdeII
    >< MboI
    >< DpnII
    > < DpnI
    FIGURA 13. 58

```

```

                >< BspAI
                > < Bsp143I
                >< BsgI                >< AlwI                >< BsrI                BspWI >
GATTTGTTTA TGAGATTTTT TACTCTTGGA TCAATTACTG CACAGCCAGT AAAAATTGAC AATGCTTCTC
25280      25290      25300      25310      25320      25330      25340

    >< ScaI
    >< RsaI
    >< Csp6I    >< SfcI
    >< AfaI    >< NlaIII    >< AciI                >< MnlI                FokI >
CTGCAAGTAC TGTTCATGCT ACAGCAACGA TACCGCTACA AGCCTCACTC CCTTTCGGAT GGCTTGTTAT
25350      25360      25370      25380      25390      25400      25410

                > < HinPII
                > < Hin6I
                >< HhaI                RmaI ><
                >< HaeII                >< HinPII                NheI ><
                >< Eco47III                >< Hin6I                MaeI ><
                >< CfoI                >< HhaI                Fnu4HI ><
                >< BspWI                >< Bsp143II                >< CfoI                AluI ><
TGGCGTTGCA TTTCTTGCTG TTTTTCAGAG CGCTACCAAA ATAATTGCGC TCAATAAAAAG ATGGCAGCTA
25420      25430      25440      25450      25460      25470      25480

    >< EcoNI
    >< BslI
    >< BsiYI                >< MaeIII
    >< BbvI    >< BsrI    >< BbvI    > < Fnu4HI                BbvI ><
GCCCTTTATA AGGGCTTCCA GTTCATTTGC AATTTACTGC TGCTATTTGT TACCATCTAT TCACATCTTT
25490      25500      25510      25520      25530      25540      25550

                Zsp2I ><
                Ppu10I ><
                > < SfcI                >< HinPII                NsiI ><
                >< PstI                >< Hin6I                >< RsaI                Mph1103I ><
                > < Fnu4HI                >< HhaI                >< Csp6I                EcoT22I ><
    >< BspMI    >< MnlI                >< CfoI                >< AfaI                >< MnlI                AvaIII ><
TGCTTGTCGC TGCAGGTATG GAGGCGCAAT TTTTGACCT CTATGCCTTG ATATATTTTC TACAATGCAT
25560      25570      25580      25590      25600      25610      25620

    >< SfaNI
    >< NspI
    >< NspHI
    >< NlaIII                >< SfaNI
CAACGCATGT AGAATTATTA TGAGATGTTG GCTTTGTTGG AAGTGCAAAT CCAAGAACCC ATTACTTTAT
25630      25640      25650      25660      25670      25680      25690

                >< Bst1107I
                >< AccI                MaeIII ><
GATGCCAACT ACTTTGTTTG CTGGCACACA CATAACTATG ACTACTGTAT ACCATATAAC AGTGTACAG
25700      25710      25720      25730      25740      25750      25760

                >< MboII
                BstXI ><
                >< HphI
    >< MunI >< MaeIII    >< MaeIII                >< Eco57I                >< BbsI MnlI >
ATACAATTGT CGTTACTGAA GGTGACGGCA TTTCAACACC AAAACTCAA GAAGACTACC AAATTGGTGG
25770      25780      25790      25800      25810      25820      25830

                >< RsaI
                > < NlaIII
                >< HphI
                >< Tru9I >< Tth111I>< Csp6I
    >< DdeI                >< DdeI    >< MseI>< AspI                >< AfaI

```

FIGURA 13. 59

```

TTATTCTGAG GATAGGCACT CAGGTGTTAA AGACTATGTC GTTGTACATG GCTATTTTCA CGAAGTTTAC
25840      25850      25860      25870      25880      25890      25900

                > < HinfI>< P1eI                >< BsrI                Tru9I ><
                >< AluI >< AccI                >< SfcI >< AlwNI                MseI ><
TACCAGCTTG AGTCTACACA AATTACTACA GACACTGGTA TTGAAAATGC TACATTCTTC ATCTTTAACA
25910      25920      25930      25940      25950      25960      25970

                > < TthHB8I
                >< Tru9I                > < TaqI                >< Ksp632I
                >< MseI                > < MboII                >< EarI BspWI ><
>< AluI                >< Eco57I                >< Eam1104I AlwI ><
AGCTTGTTAA AGACCCACCG AATGTGCAAA TACACACAAT CGACGGCTCT TCAGGAGTTG CTAATCCAGC
25980      25990      26000      26010      26020      26030      26040

>< XhoII
>< Sau3AI
>< NlaIV
>< NdeII
>< MflI
>< MboI
>< DpnII
>< DpnI
>< BstYI
>< BstI
>< BspAI
>< Bsp143I
>< BscBI
>< BamHI >< AlwI
AATGGATCCA ATTTATGATG AGCCGACGAC GACTACTAGC GTGCCTTTGT AAGCACAAGA AAGTGAGTAC
26050      26060      26070      26080      26090      26100      26110

                > < Tru9I
                >< RsaI
                > < MseI
                >< MboII
                > < RsaI                >< MaeII                >< RsaI
                >< Csp6I                >< Csp6I                >< Tru9I >< Csp6I
                > < AfaI                >< AfaI                >< MseI >< AfaI
GAAGTTATGT ACTCATTCGT TTCGGAAGAA ACAGGTACGT TAATAGTTAA TAGCGTACTT CTTTTCTTGT
26120      26130      26140      26150      26160      26170      26180

                >< TthHB8I
                >< TaqI
                >< RmaI                >< HinfII                > < RsaI
                > < MaeIII                >< HinfI                Fnu4HI ><
                >< MaeI >< RmaI                >< HhaI                >< Csp6I
                >< FokI >< MaeI                >< CfoI >< BbvI > < AfaI
CTTTCGTGGT ATTCTTGCTA GTCACACTAG CCATCCTTAC TGCGCTTCGA TTGTGTGCGT ACTGCTGCAA
26190      26200      26210      26220      26230      26240      26250

                >< Tru9I
                >< Tru9I                >< ThaI
                >< MseI                >< MvnI
                >< SspI >< MaeII                >< MseI
                >< HpaI                >< BstUI                Ksp632I >
                >< HindII                >< MaeII >< Bsp50I >< MboII EarI >
                >< HincII                >< AccI >< AccII                Eam1104I >
TATTGTAAAC GTGAGTTTAG TAAAACCAAC GGTTCACGTC TACTCGCGTG TTAAAAATCT GAAGTCTTCT
26260      26270      26280      26290      26300      26310      26320

```

FIGURA 13. 60

```

>< Sau3AI
>< NdeII
>< MboI
>< DpnII
>< MboII>< DpnI
>< XmnI >< BspAI> < Eco57I
>< Asp700I>< BspI43I
GAAGGAGTTC CTGATCTTCT GGTCTAAACG AACTAACTAT TATTATTATT CTGTTTGGAA CTTTAACATT
26330      26340      26350      26360      26370      26380      26390

>< ScrFI
>< MvaI
>< EcoRII
>< Ecl136I
>< DsaV NlaIV ><
>< RsaI
>< MnlI
>< Tru9I
>< Csp6I
>< MseI
>< BstNI RmaI ><
> < NlaIII >< AfaI > < AluI >< ApyIBscBI ><
GCTTATCATG GCAGACAACG GTACTATTAC CGTTGAGGAG CTAAACAAC TCCTGGAACA ATGGAACCTA
26400      26410      26420      26430      26440      26450      26460

>< ScrFI
>< RmaI
>< MvaI
>< MaeI
>< EcoRII
>< Ecl136I
>< DsaV
>< BstOI
>< BstNI
>< BsiLI
>< ApyI >< MaeIII
GTAATAGGTT TCCTATTCCT AGCCTGGATT ATGTTACTAC AATTTGCCTA TTCTAATCGG AACAGGTTTT
26470      26480      26490      26500      26510      26520      26530

>< Pali
>< MscI
>< MnlI >< MaeIII
>< HaeIII
>< EaeI
>< BsuRI
>< BsrI
>< RsaI
>< Csp6I >< HindIII
>< AfaI >< AluI
>< BspWI
>< BshI
>< Bali >< BbvI Fnu4HI ><
TGTAATAAT AAAGCTTGTT TTCCTCTGGC TCTTGTTGGCC AGTAACACTT GCTTGTTTTG TGCTTGCTGC
26540      26550      26560      26570      26580      26590      26600

>< VspI
>< Tru9I
>< MseI
>< HphI
>< SfcI >< AsnI >< BsrI
>< AccI >< AseI>< MaeIII>< AciI
TGTCTACAGA ATTAATTGGG TGACTGGCGG GATTGCGATT GCAATGGCTT GTATTGTAGG CTTGATGTGG
26610      26620      26630      26640      26650      26660      26670

>< EspI
>< Eco57I
>< DdeI
>< CelII
>< Bpu1102I
>< RsaI
>< Csp6I

```

FIGURA 13. 61

```

>< BfrI                                     >< AfaI
  >< AluI                                     >< AciI                                     MboII >
CTTAGCTACT TCGTTGCTTC CTCAGGCTG TTTGCTCGTA CCCGCTCAAT GTGGTCATTC AACCCAGAAA
  26680      26690      26700      26710      26720      26730      26740

  >< ScrFI
  >< NciI
  >< MspI
  >< HpaII
  >< HapII
  >< DsaV>< MnlI
  >< BslI
  >< BsiYI
  >< BsaJI >< MunI                                     > < XcmI
  >< BcnI                                     >< MaeIII >< AciI >< NlaIII
CAAACATTCT TCTCAATGTG CCTCTCCGGG GGACAATTGT GACCAGACCG CTCATGGAAA GTGAACTTGT
  26750      26760      26770      26780      26790      26800      26810

Tru9I ><
  SinI >
  Sau96I >
  PpuMI >
  NspIV >
  MseI ><
  >< MaeIII
  >< Sau3AI                                     > < RmaI >< HaeII
  >< NdeII                                     >< PstI                                     > < MaeI   EcoO109I >
  >< MboI                                     >< MspI                                     >< HinPIIEco47I >
    >< FbaI                                     >< HpaII >< StyI>< Hin6I   DraII >
  >< DpnII                                     >< HapII >< EcoT14I   Cfr13I >
    >< DpnI                                     >< HaeIII >< EcoI30I>< Bsp143II
  >< BspAI                                     >< GdiI                                     >< BssT1I   Bsi2I >
    >< Bsp143I                                     >< EaeI                                     >< BsaJI   Bme18I >
  >< BsiQI                                     >< BsuRI                                     >< BlnI >< HhaI   AvaII >
  >< BclI                                     >< MaeIII                                     >< AvrII >< CfoI   AsuI >
CATTGGTGCT GTGATCATTC GTGGTCACTT GCGAATGGCC GGACACTCCC TAGGGCGCTG TGACATTAAG
  26820      26830      26840      26850      26860      26870      26880

  >< Sau3AI
  >< NdeII
  >< MboI
  >< DpnII
    >< DpnI
  >< PssI >< BspMI
  >< Psp5II >< BspAI                                     >< XmnI
  >< NspHII >< Bsp143I                                     >< Asp700I > < HgaI   Fnu4HI ><
GACCTGCCAA AAGAGATCAC TGTGGCTACA TCACGAACGC TTTCTTATTA CAAATTAGGA GCGTCGCAGC
  26890      26900      26910      26920      26930      26940      26950

  >< TfiI
  >< HinfI
  >< BbvI                                     > < Tru9I
  >< BbvI                                     >< Fnu4HI >< AciI                                     > < MseI
GTGTAGGCAC TGATTGAGGT TTTGCTGCAT ACAACCGCTA CCGTATTGGA AACTATAAAT TAAATACAGA
  26960      26970      26980      26990      27000      27010      27020

  >< MspI                                     >< RsaI
  >< HpaII                                     >< RmaI
  >< HapII                                     >< Csp6I
  >< Cfr10I                                     >< MaeI>< BcgI                                     HindII ><
  >< BcgI/a                                     >< SspI                                     >< AfaI >< MaeIII   HincII ><

```

FIGURA 13. 62

```

CCACGCCGGT AGCAACGACA ATATTGCTTT GCTAGTACAG TAAGTGACAA CAGATGTTTC ATCTTGTTGA
 27030      27040      27050      27060      27070      27080      27090

>< ScrFI
>< MvaI
  >< MaeIII
>< EcoRII
  >< Ecl136I
>< DsaV
  >< BstOI
  >< BstNI
  >< BsiLI
  >< ApyI
                                >< MnlI
                                HinfI ><
CTTCCAGGTT ACAATAGCAG AGATATTGAT TATCATTATG AGGACTTTCA GGATTGCTAT TTGGAATCTT
 27100      27110      27120      27130      27140      27150      27160

                                >< BsmAI
                                >< Tru9I
                                > < MnlI
>< MaeII
  >< Alw26I
                                >< MseI
                                >< DdeI
                                >< MboII
GACGTTATAA TAAGTTCAAT AGTGAGACAA TTATTTAAGC CTCTAACTAA GAAGAATTAT TCGGAGTTAG
 27170      27180      27190      27200      27210      27220      27230

                                >< Ksp632I
                                >< MboII
                                >< EarI
                                >< NlaIII Eam1104I ><
ATGATGAAGA ACCTATGGAG TTAGATTATC CATAAAACGA ACATGAAAAT TATTCTCTTC CTGACATTGA
 27240      27250      27260      27270      27280      27290      27300

                                > < RsaI >< RsaI
                                >< Csp6I >< Csp6I
                                > < AluI
                                >< MnlI
                                > < AfaI >< AfaI
TTGTATTTAC ATCTTGCGAG CTATATCACT ATCAGGAGTG TGTTAGAGGT ACGACTGTAC TACTAAAAGA
 27310      27320      27330      27340      27350      27360      27370

                                >< MnlI >< HphI >< HphI
                                >< MnlI
ACCTTGCCCA TCAGGAACAT ACGAGGGCAA TTCACCATT CACCCTCTTG CTGACAATAA ATTTGCACTA
 27380      27390      27400      27410      27420      27430      27440

                                Sau3AI >
                                > < PvuII
                                > < Psp5I
                                > < NspBII
                                >< TthHB8I
                                >< TaqI
                                >< RsaI
                                >< Csp6I
                                >< BbvI
                                >< AfaI
                                > < AluI
                                >< RmaI
                                >< MaeI
ACTTGCACTA GCACACACTT TGCTTTTGCT TGTGCTGACG GTACTCGACA TACCTATCAG CTGCGTGCAA
 27450      27460      27470      27480      27490      27500      27510

                                >< SstI
                                >< SduI
                                >< SacI
                                >< NspII
                                >< HgiAI
                                >< Eco24I
                                > < Ecl136II
                                >< BspWI
                                >< Bsp1286I
                                >< BmyI
                                >< BanII
                                >< Alw21I
>< HphI
>< DpnI
                                >< MnlI

```

FIGURA 13. 63

ES 2 396 127 T3

```

>< Bsp143I          >< MnlI          > < AluI      BbvI ><
GATCAGTTTC ACCAAACTT TTCATCAGAC AAGAGGAGGT TCAACAAGAG CTCTACTCGC CACTTTTTCT
27520      27530      27540      27550      27560      27570      27580

SstI ><
SduI ><
SacI ><
NspII ><
HgiAI ><
Eco24I ><
Ecl136II ><
Bsp1286I ><
BmyI ><
BanII ><
Alw21I ><
AluI ><
>< Fnu4HI          >< HphI          >< Tru9I          >< Tru9I
>< MaeI          >< MseI          >< MseI          >< MseI
CATTGTTGCT GCTCTAGTAT TTTTAATACT TTGCTTCACC ATTAAGAGAA AGACAGAATG AATGAGCTCA
27590      27600      27610      27620      27630      27640      27650

>< Tru9I          >< Tru9I
>< MseI          >< MseI
CTTTAATTGA CTTCTATTTG TGCTTTTTAG CCTTTCTGCT ATTCCTTGTT TTAATAATGC TTATTATATT
27660      27670      27680      27690      27700      27710      27720

>< XhoII
>< XbaI
> < ScrFI
>< Sau3AI
>< RmaI
>< NdeII
> < MvaI
>< MflI
>< MboI
>< EcoRII>< MaeI
> < Ecl136I
>< DpnII
>< DpnI
>< BstYI
> < BstOI
> < BstNI
>< TthHB8I >< BspAI          > < RsaI
>< DsaV>< Bsp143I          >< MboII
> < BsiLI          >< Csp6I
>< TaqI > < ApyI > < AlwI > < AfaI          >< NlaIII
TTGGTTTTCA CTCGAAATCC AGGATCTAGA AGAACCTTGT ACCAAAGTCT AAACGAACAT GAAACTTCTC
27730      27740      27750      27760      27770      27780      27790

>< HinPII
>< Hin6I
>< HhaI
>< RsaI >< HaeII
>< SfcI          >< Eco47III
>< Csp6I>< CfoI SfaNI ><
>< AfaI >< Bsp143II
ATTGTTTTGA CTTGTATTTT TCTATGCAGT TGCATATGCA CTGTAGTACA GCGCTGTGCA TCTAATAAAC
27800      27810      27820      27830      27840      27850      27860

>< XhoII
>< Sau3AI
>< NdeII
> < MnlI
>< MflI

```

FIGURA 13. 64

```

>< MboI
>< DpnII
>< DpnI >< RsaI
>< BstYI >< MboII
>< NlaIII>< BspAI >< Csp6I >< RmaI
>< AlwI >< Bsp143I >< AfaI >< MaeI
CTCATGTGCT TGAAGATCCT TGTAAGGTAC AACACTAGGG GTAATACTTA TAGCACTGCT TGGCTTTGTG
27870 27880 27890 27900 27910 27920 27930

>< SduI
>< RmaI
>< NspII
>< MaeI
>< HgiAI
>< Bsp1286I >< NspI
>< BmyI >< NspHI
>< Alw21I >< NlaIII >< MaeIII
CTCTAGGAAA GGTTTTACCT TTTCATAGAT GGCACACTAT GGTTCAAACA TGCACACCTA ATGTTACTAT
27940 27950 27960 27970 27980 27990 28000

> < XhoII
> < Sau3AI > < Van91I >< RsaI
>> PvuII >< NlaIV
>> Psp5I >< KpnI >< NlaIII
> < NdeII > < PflMI >< Eco64I >< MaeIII
> < MflI>< NspBII >< Csp6I>< HphI
> < DpnII >< HinPII >< BscBI >< EcoO65I
>> Bsp143I >< Hin6I >< Bani >< BspHI
> < BstYI > < BslI >< HhaI >< RmaI >< Asp718 >< Eco91I
> < BspAI > < BsiYI>< CfoI >< MaeI >< AfaI >< BstPI
> < MboI>< AluI>< BspWI >< BspWI >< AccB1I >< BstEII
>< AlwI >< DpnI > < AccB7I >< AluI >< Acc65I >< BbvI
CAACTGTCAA GATCCAGCTG GTGGTGCCT TATAGCTAGG TGTTGGTACC TTCATGAAGG TCACCAAAC
28010 28020 28030 28040 28050 28060 28070

>< SinI
>< Sau96I
>< NspIV
NspHII ><
NlaIV ><
>< Eco47I
>< Cfr13I
>< BsiZI
BscBI ><
>< Esp3I >< Csp6I >< Tru9I >< Bmel8I
>< BsmAI >< BsmBI >< MseI >< Tru9I >< AvaII
>< Alw26I >< AfaI >< DraI >< MseI >< AsuI
GCTGCATTTA GAGACGTA CTGTTGTTTAA AATAAACGAA CAAATTAAAA TGTCTGATAA TGGACCCCAA
28080 28090 28100 28110 28120 28130 28140

>< SinI
>< Sau96I
>< NspIV
>< NspHII
>< NlaIV
>< Eco47I
>< Cfr13I
>< BsiZI
>< SduI >< BscBI
>< NspII >< Bmel8I
>< Bsp1286I >< AvaII >< TfiI
>< BmyI >< AsuI >< HinfI
>< MaeII >< AciI >> MnlI

```

FIGURA 13. 65

```

TCAAACCAAC GTAGTGCCCC CCGCATTACA TTTGGTGGAC CCACAGATTC AACTGACAAT AACCAGAATG
28150      28160      28170      28180      28190      28200      28210

      >< HinPII >< StyI
      >< HaeII
      > < Pali >< Hin6I >< EcoT14I
      > < HaeIII >< HhaI>< Eco130I
      >< BspWI >< BssTII
      > < BsuRI >< Bsp143II
      >< HgaI> < BshI >< CfoI>< BsaJI >< HgaI
GAGGACGCAA TGGGGCAAGG CCAAAACAGC GCCGACCCCA AGGTTTACCC AATAATACTG CGTCTTGGTT
28220      28230      28240      28250      28260      28270      28280

      >< TthHB8I
      > < ScrFI
      >< Pali
      >< PaeR7I
      >< NspIII
      > < MvaI
      >< HaeIII
      >< EcoRII
      >< Eco88I
      >< XhoI > < Ecl136I
      >< DsaV
      >< BsuRI
      >< SlaI > < BstOI
      >< MnlI>< TaqI> < BstNI
      >< CcrI > < BsiLI
      >< HinfI >< BshI
      >< TfiI>< BcoI>< BsaJI
      >< MnlI >< DdeI >< Aval > < ApyI
      >< AluI >< DdeI > < NlaIII >< BfrI >< Ama87I >< MnlI
CACAGCTCTC ACTCAGCATG GCAAGGAGGA ACTTAGATTC CCTCGAGGCC AGGGCGTTCC AATCAACACC
28290      28300      28310      28320      28330      28340      28350

      >< SinI
      >< Sau96I
      >< NspIV
      >< NspHII
      >< Eco47I
      >< Cfr13I
      >< BsiZI
      >< Bme18I > < Ksp632I
      >< Aval > < Eam1104I
      >< AsuI > < EarI > < AluI>< MboII >< MaeIII
AATAGTGGTC CAGATGACCA AATTGGCTAC TACCGAAGAG CTACCCGACG AGTTCGTGGT GGTGACGGCA
28360      28370      28380      28390      28400      28410      28420

      >< SstI
      >< SduI
      >< SacI
      >< NspII
      >< HgiAI
      >< EspI
      >< Eco24I
      >< Ecl136II >< StyI >< Sau96I
      >< DdeI >< RmaI >< NspIV
      >< CelII >< MaeI >< HaeIII
      >< Bsp1286I >< EcoT14I >< Cfr13I
      >< Bpu1102I >< Eco130I >< BsuRI
      >< BmyI >< BssTII > < BsrI
      >< BanII >< RsaI >< BsaJI >< BsiZI

```

FIGURA 13. 66

```

>< Alw21I    >< Csp6I    >< BlnI    >< BshI>< HindIII
>< HphI >< AluI    >< AfaI    >< AvrII    >< AsuI    >< AluI
AAATGAAAGA GCTCAGCCCC AGATGGTACT TCTATTACCT AGGAACTGGC CCAGAAGCTT CACTTCCCTA
28430      28440      28450      28460      28470      28480      28490

>< HinPII
>< Hin6I
>< HhaI
>< HaeII
>< CfoI          > < MnlI    >< NlaIV
>< Bsp143II      >< SfaNI >< DdeI >< BscBI
CGGCGCTAAC AAAGAAGGCA TCGTATGGGT TGCAACTGAG GGAGCCTTGA ATACACCCAA AGACCACATT
28500      28510      28520      28530      28540      28550      28560

>< NlaIV
>< Eco64I
>< BscBI
>< Bani
>< AciI
>< AccBII >< BbvI    >< Fnu4HI    >< MnlI
GGCACCCGCA ATCCTAATAA CAATGCTGCC ACCGTGCTAC AACTTCCTCA AGGAACAACA TTGCCAAAAG
28570      28580      28590      28600      28610      28620      28630

>< ThaI
>< MnlI
>< MaeII >< MvnI
>< MnlI    BstUI ><
>< Fnu4HI    >< Ksp632I    Bsp50I ><
>< BspWI    >< EarI    >< BsaAI>< AciI
>< MnlI    >< MnlI    >< AciI>< MboII    >< Eam1104I    AccII ><
GCTTCTACGC AGAGGGAAGC AGAGGCGGCA GTCAAGCCTC TTCTCGCTCC TCATCACGTA GTCGCGGTAA
28640      28650      28660      28670      28680      28690      28700

>< ScrFI
>< MvaI
>< EcoRII
>< Ecl136I
>< DsaV>< Fnu4HI
>< BstOI
>< BstNI
>< BsiLI
>< ApyI
>< BbvI
>< TaqI
>< AciI
TTCAAGAAAT TCAACTCCTG GCAGCAGTAG GGGAAATTCT CCTGCTCGAA TGGCTAGCGG AGGTGGTGAA
28710      28720      28730      28740      28750      28760      28770

> < ThaI
> < MvnI
>< HphI    >< MnlI
> < HinPII
> < Hin6I
>< HhaI
> < BstUI    >< RmaI
> < Bsp50I    >< MaeI
>< BbvI >< CfoI>< Fnu4HI
> < AccII>< BspWI
>< AluI
ACTGCCCTCG CGCTATTGCT GCTAGACAGA TTGAACAGC TTGAGAGCAA AGTTTCTGGT AAAGGCCAAC
28780      28790      28800      28810      28820      28830      28840

> < PalI>< MaeIII
> < HaeIII
> < BsuRI    >< DdeI
>< Fnu4HI
>< DdeI
>< RsaI ><
>< MnlI
>< MaeII ><
>< Csp6I ><

```

FIGURA 13. 67

FIGURA 13. 68

```

                                >< PfuI
    >< Fnu4HI
    >< BspWI
    >< BsmAI
    >< Alw26I
    >< AciI
                                >< MboII
                                >< MboII
                                >< Ksp632I >< GsuI
                                >< MaeIII >< EarI>< Fnu4HI
                                >< HinfI >< Eam1104I>< BpmI
                                >< Fnu4HI >< BbvI >< AciI >< NlaIII
AGCCTTTGCC GCAGAGACAA AAGAAGCAGC CCACTGTGAC TCTTCTTCCT GCGGCTGACA TGGATGATTT
29270      29280      29290      29300      29310      29320      29330

                                >< NlaIII >< HinfI NlaIII ><
    >< FokI >< AluI >< TfiI>< DdeI >< BspHI
CTCCAGACAA CTTCAAAATT CCATGAGTGG AGCTTCTGCT GATTCAACTC AGGCATAAAC ACTCATGATG
29340      29350      29360      29370      29380      29390      29400

                                >< MaeII >< AccI
ACCACACAAG GCAGATGGGC TATGTAAACG TTTTCGCAAT TCCGTTTACG ATACATAGTC TACTCTTGTC
29410      29420      29430      29440      29450      29460      29470

                                >< Tru9I
                                >< Tru9I
                                >< MseI
                                >< MseI
    >< XmnI >< HpaI
    >< EcoRI>< MaeIII >< HindII Tru9I ><
    >< Asp700I >< BsgI >< HincII MseI ><
CAGAATGAAT TCTCGTAACT AAACAGCACA AGTAGGTTTA GTTAACTTTA ATCTCACATA GCAATCTTTA
29480      29490      29500      29510      29520      29530      29540

                                XorII >
                                TthHB8I >
                                TaqI >
                                Sau3AI ><
                                RsaI ><
    >< ThaIPvuI >
    NdeII ><
    >< MnlI
    >< MvnIMcrI >
    MboI ><
    DpnII ><
    DpnI ><
    Csp6I ><
    >< BstUI
    >< HaeIII BspCI >
    BspAI ><
    >< TthHB8I >< Bsp50I
    >< Pali Bsp143I ><
    >< BsuRI BsiEI >
    >< BshIAfaI ><

    >< MnlI >< TaqI >< AciI
    >< MaeIII >< MnlI >< AccII
ATCAATGTGT AACATTAGGG AGGACTTGAA AGAGCCACCA CATTTTCATC GAGGCCACGC GGAGTACGAT
29550      29560      29570      29580      29590      29600      29610

                                >< SduI
                                >< NspII
                                >< MboII >< VspI
                                >< Ksp632I >< Eco24I >< Tru9I
    >< RsaI >< RmaI >< Fnu4HI >< Bsp1286I >< MseI
    >< Csp6I >< MaeI >< EarI >< BmyI >< AsnI
    >< AfaI >< BbvI >< AluI>< Eam1104I >< BanII >< AseI

```

FIGURA 13. 69

ES 2 396 127 T3

```

CGAGGGTACA GTGAATAATG CTAGGGAGAG CTGCCTATAT GGAAGAGCCC TAATGTGTAA AATTAATTTT
 29620      29630      29640      29650      29660      29670      29680

                >< Tru9I    >< DdeI
                >< MseI    >< BfrI
                >< NlaIII   > < AluI
AGTAGTGCTA TCCCCATGTG ATTTTAATAG CTTCTTAGGA GAATGACAAA AAAAAAAAAA AAAAAA
 29690      29700      29710      29720      29730      29740

```

FIGURA 13. 70

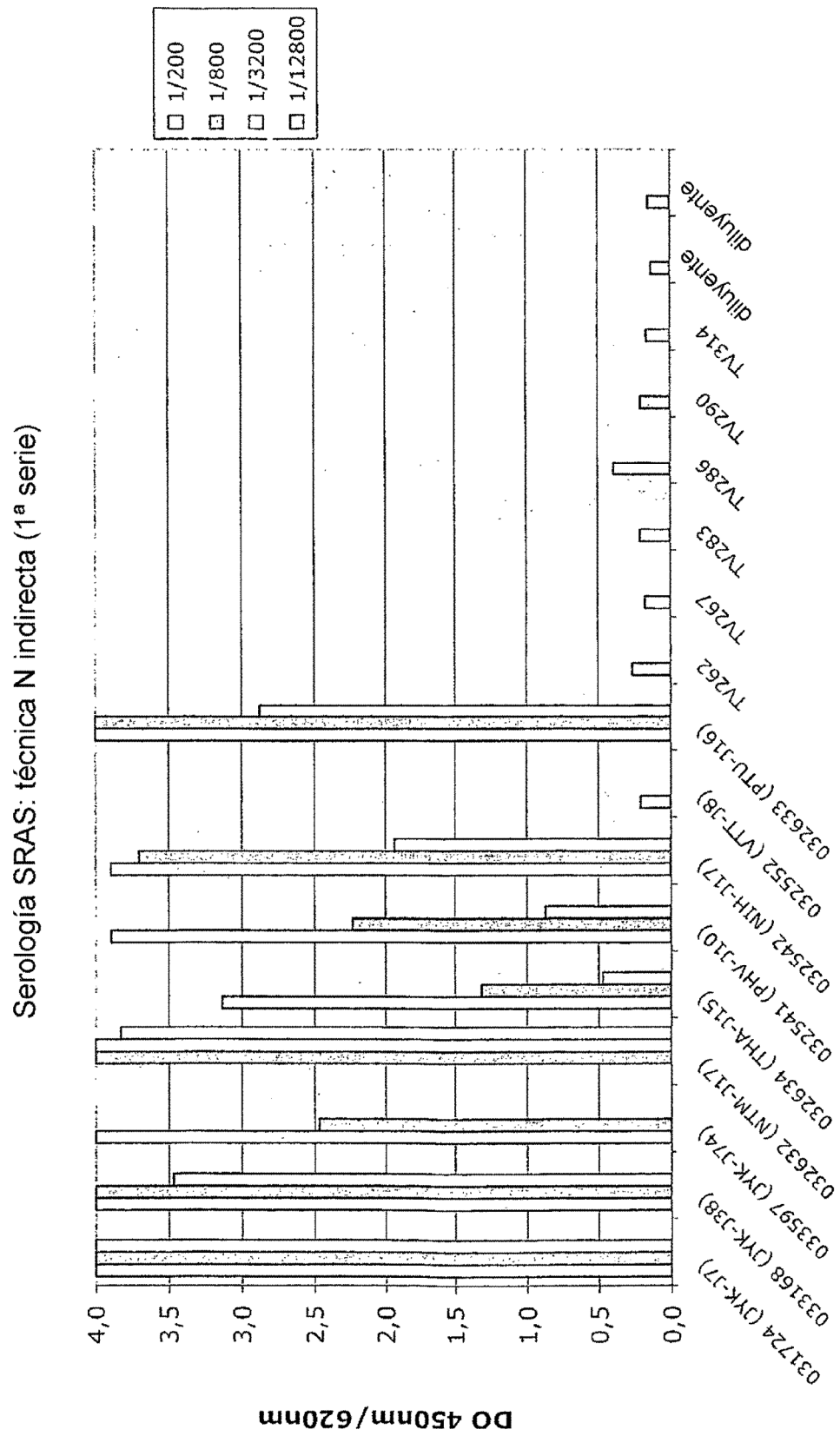


FIGURA 14

Serología SRAS: técnica N indirecta (2ª serie)

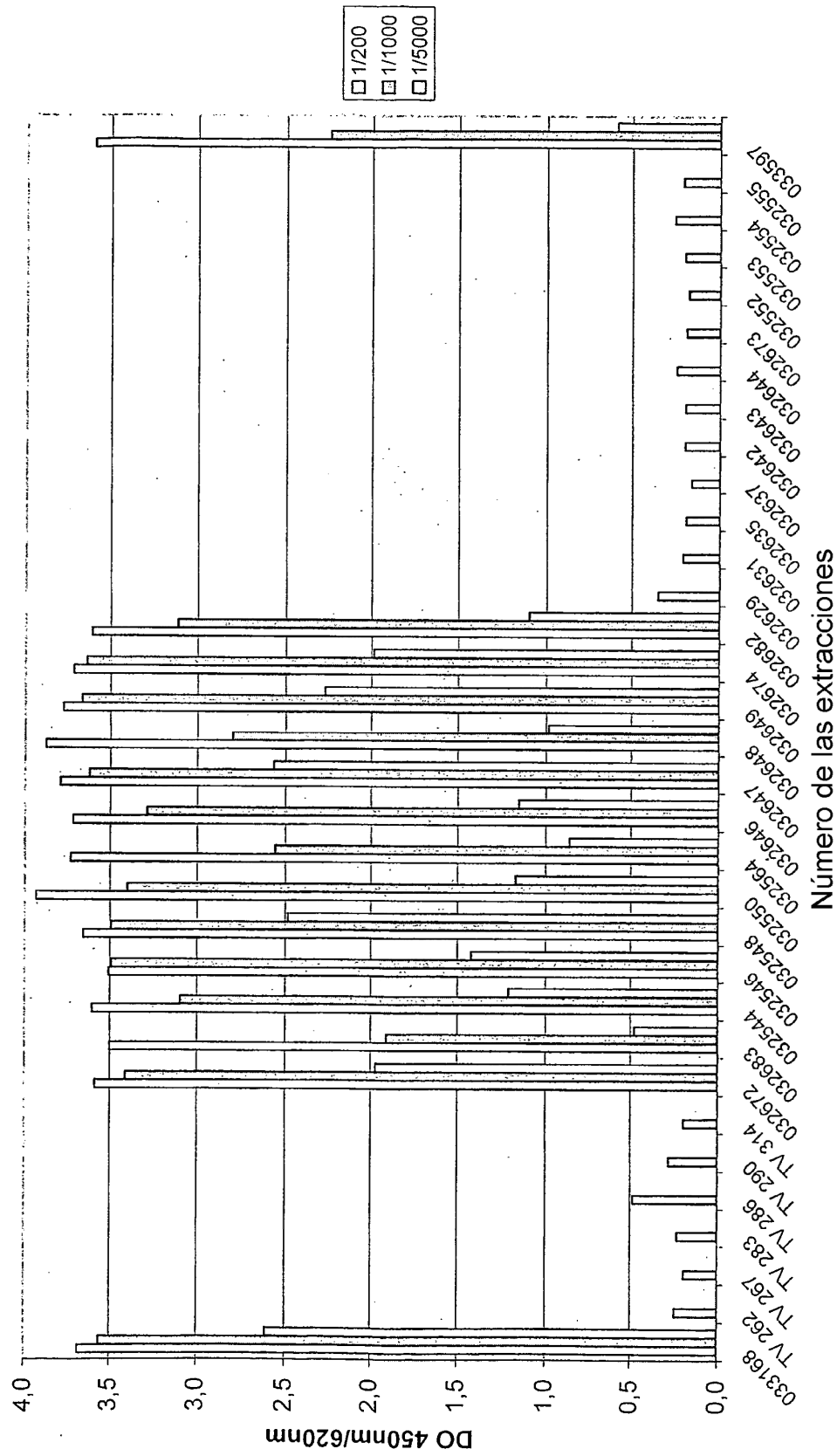


FIGURA 15

Serología SRAS: técnica de doble epítipo (1ª serie)

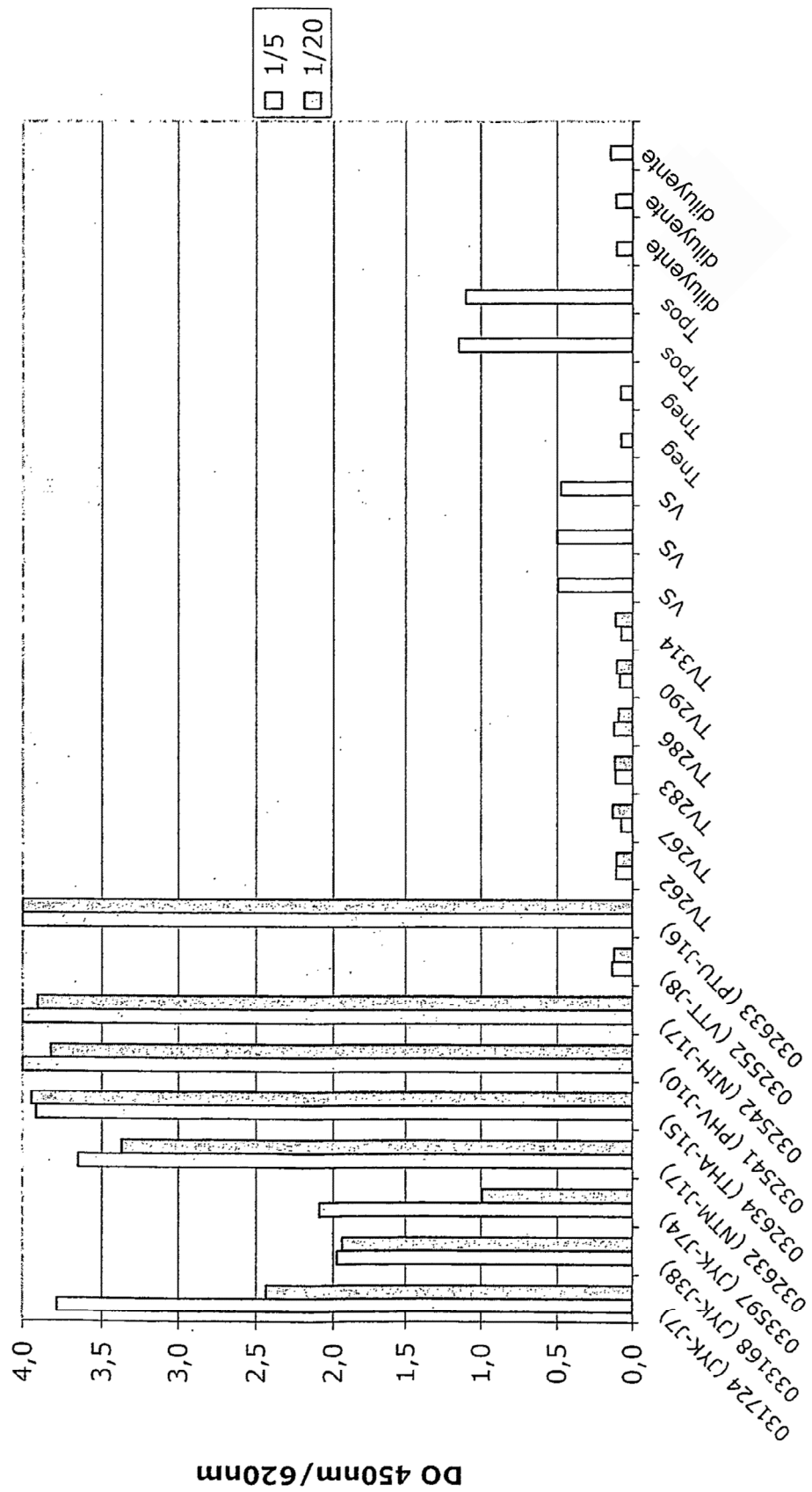


FIGURA 16

Serología SRAS: técnica de doble epítipo (2ª serie)

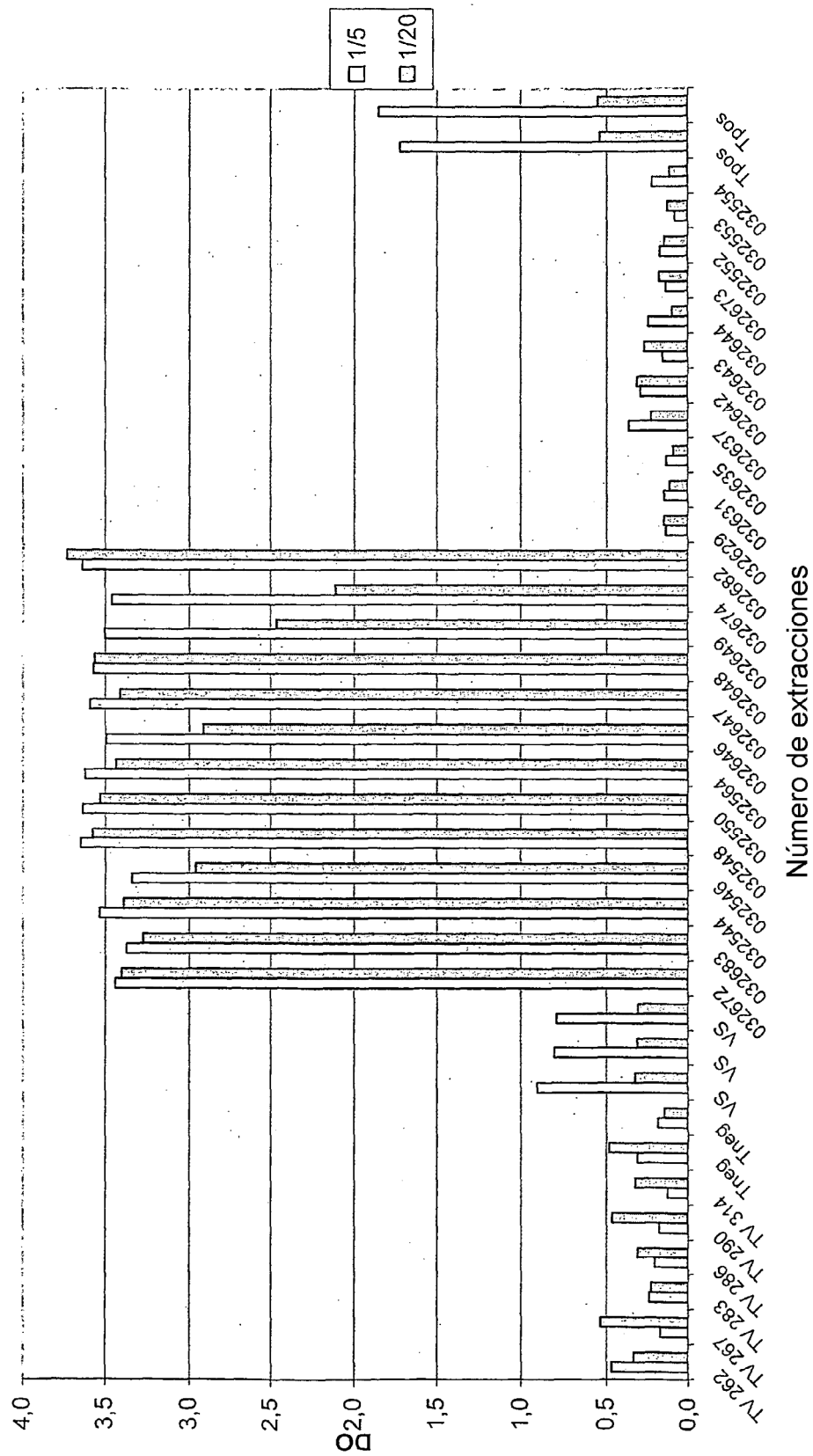


FIGURA 17

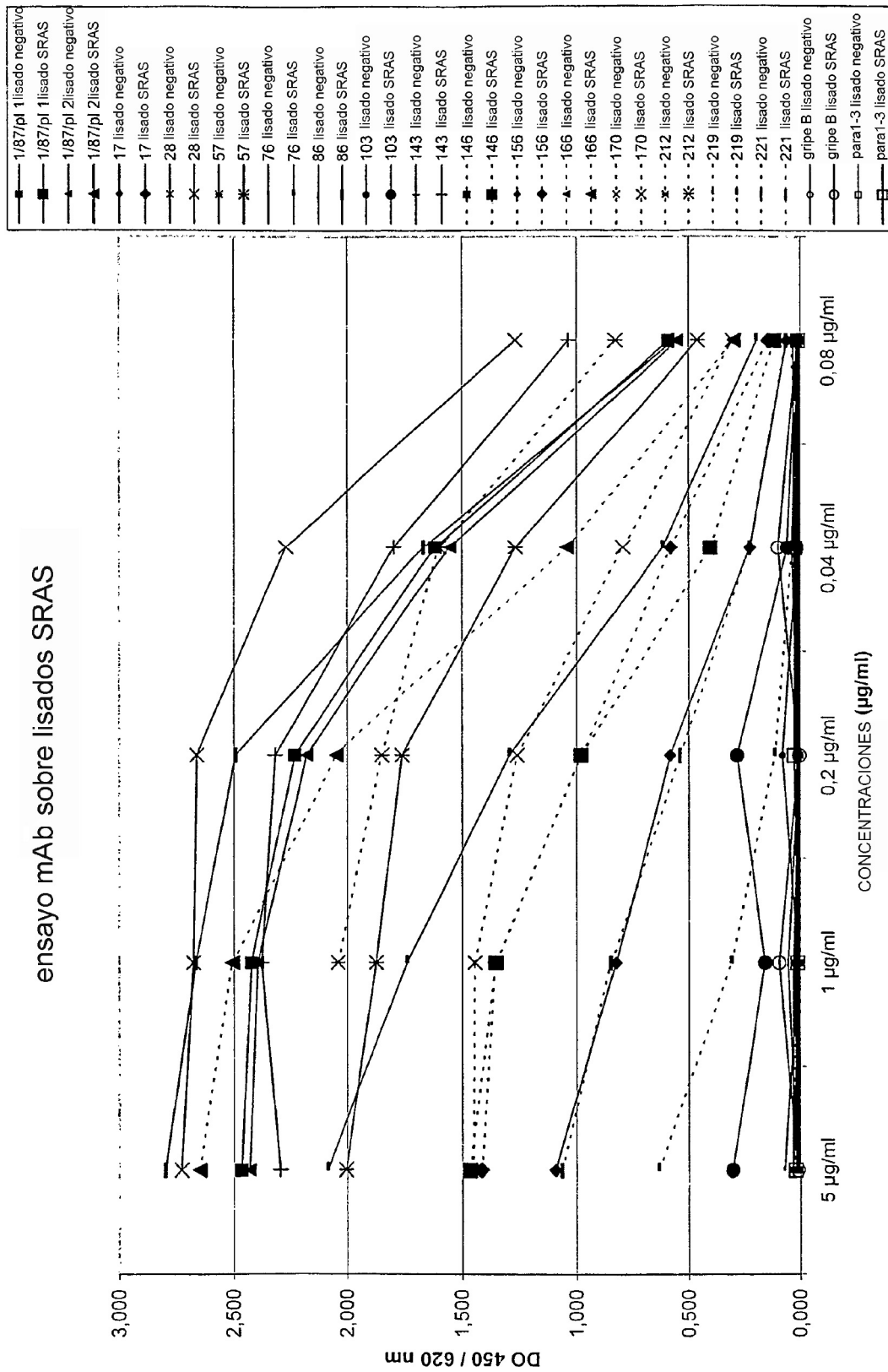


FIGURA 18

Ensayo mAb sobre lisados HCoV-229E

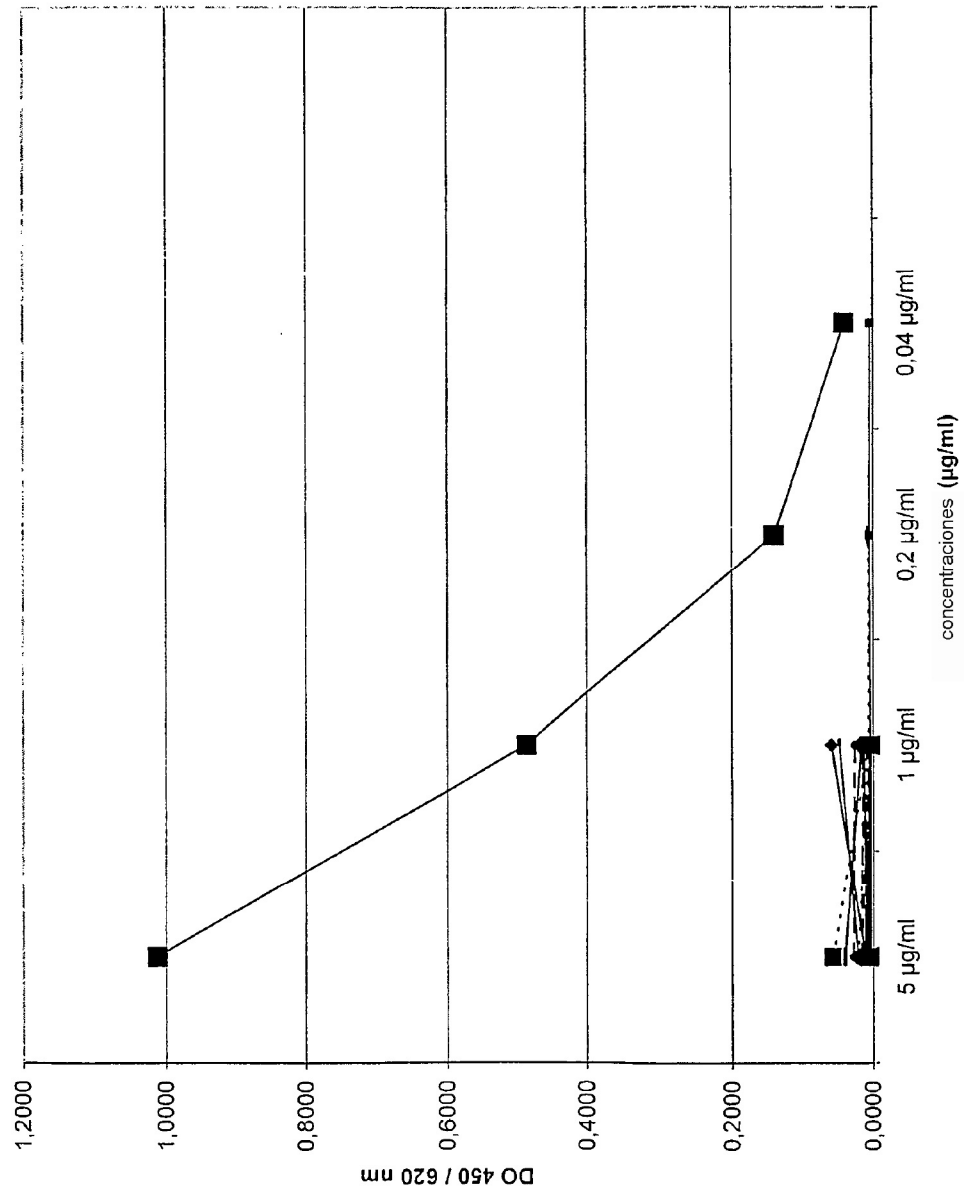
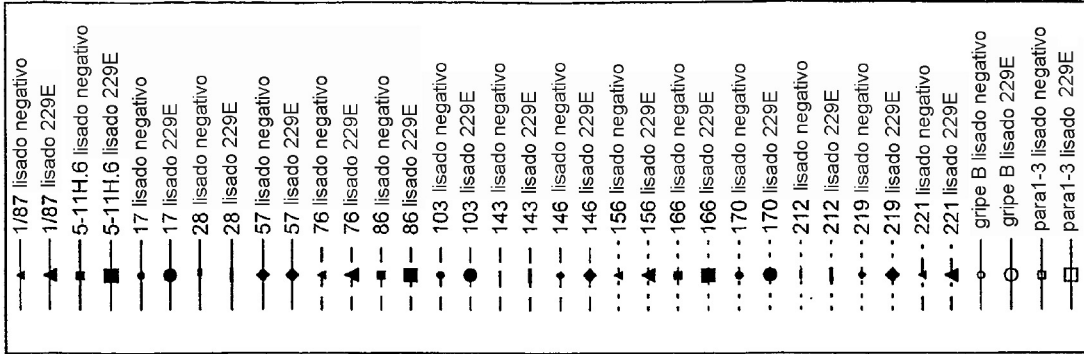


FIGURA 19



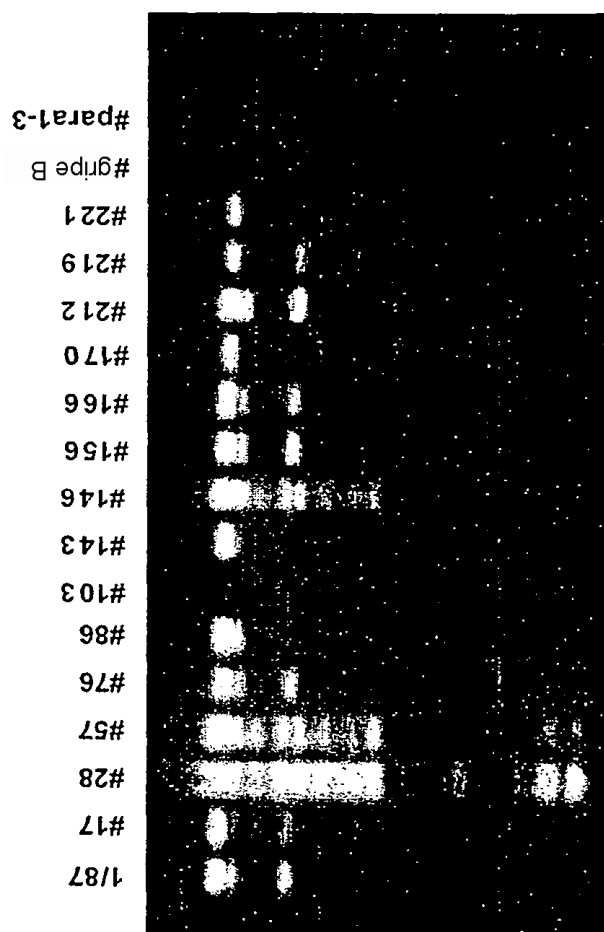


FIGURA 20

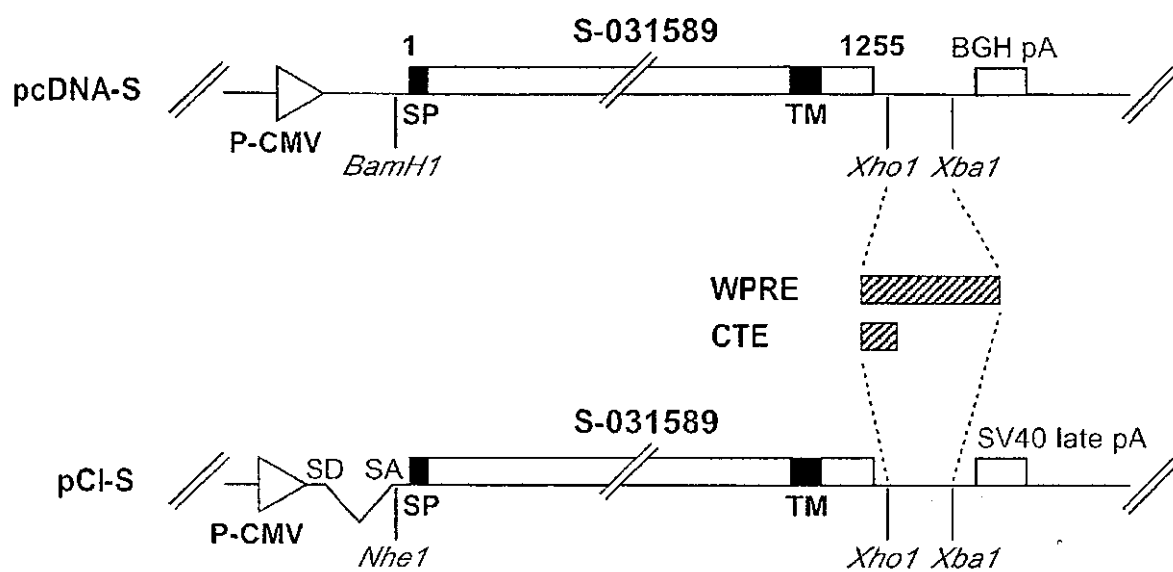


FIGURA 21

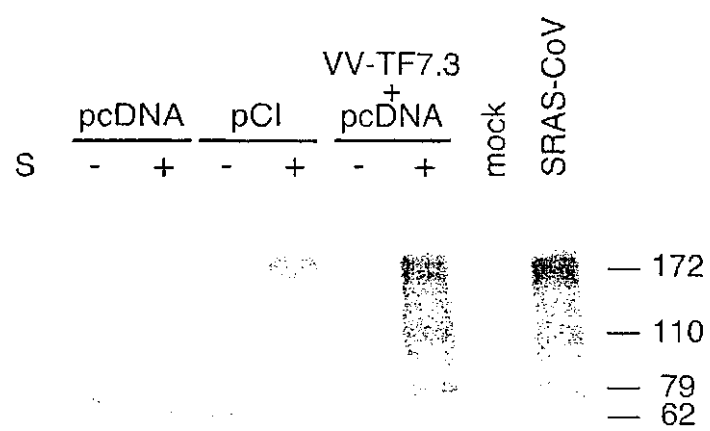
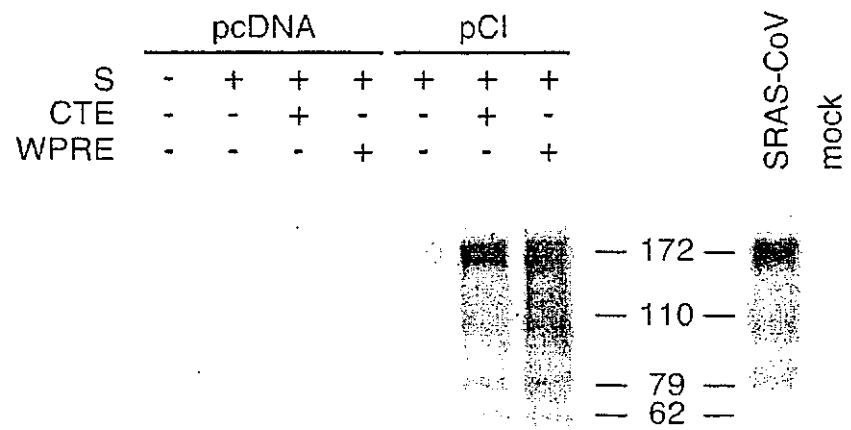


FIGURA 22

A.



B.

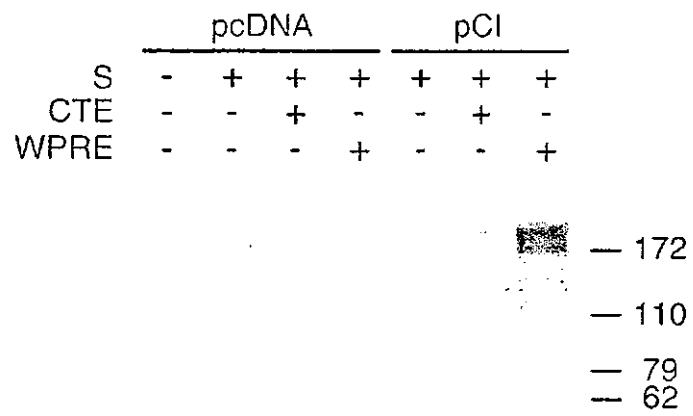
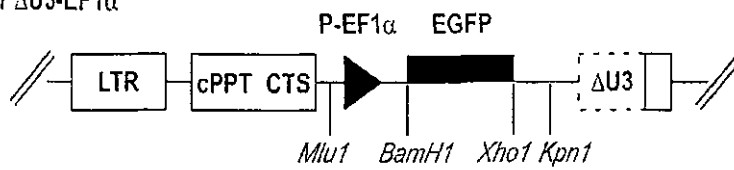
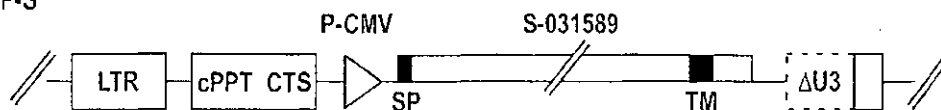


FIGURA 23

pTRIP Δ U3-EF1 α



pTRIP-S



pTRIP-SD/SA-S-CTE



pTRIP-SD/SA-S-WPRE

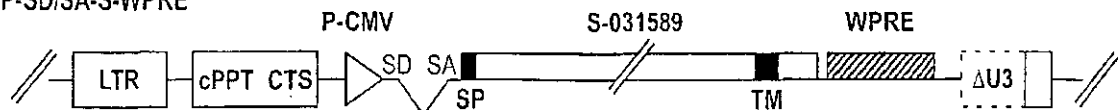


FIGURA 24

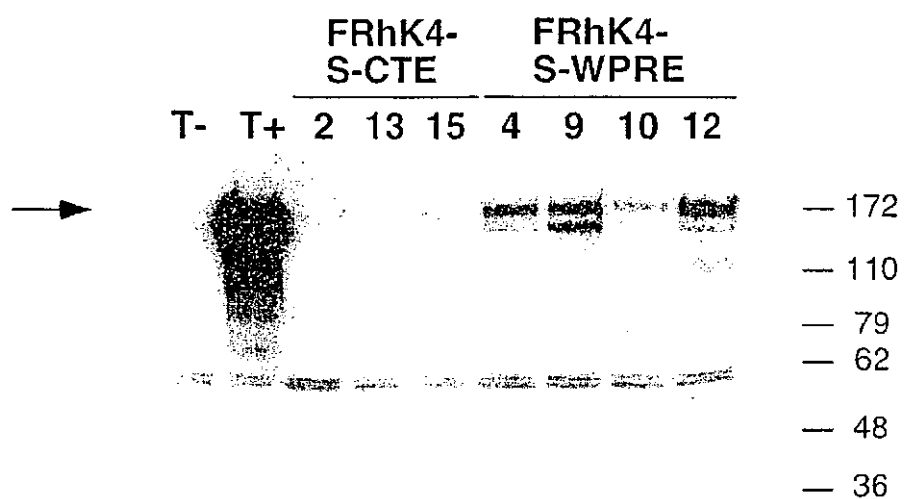


FIGURA 25

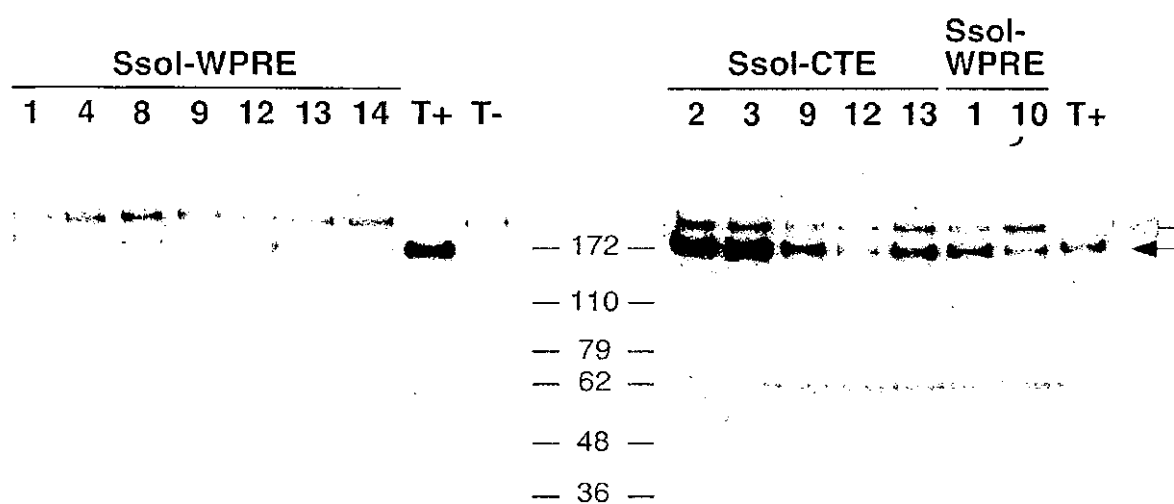
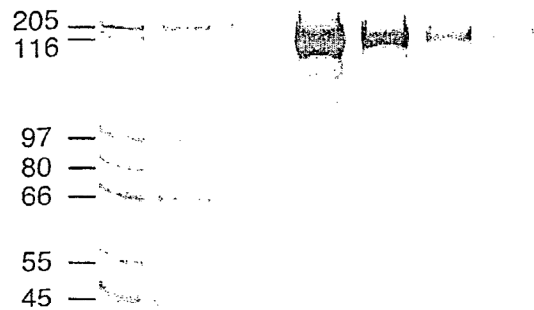


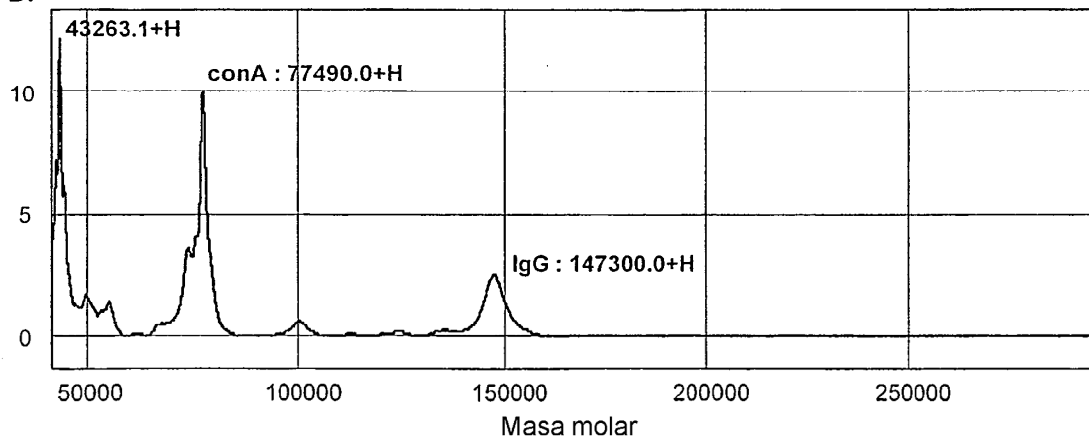
FIGURA 26

A.

MM (ng)			Ssol (μg)			
120	60	30	2	.5	.13	.03



B.



C.

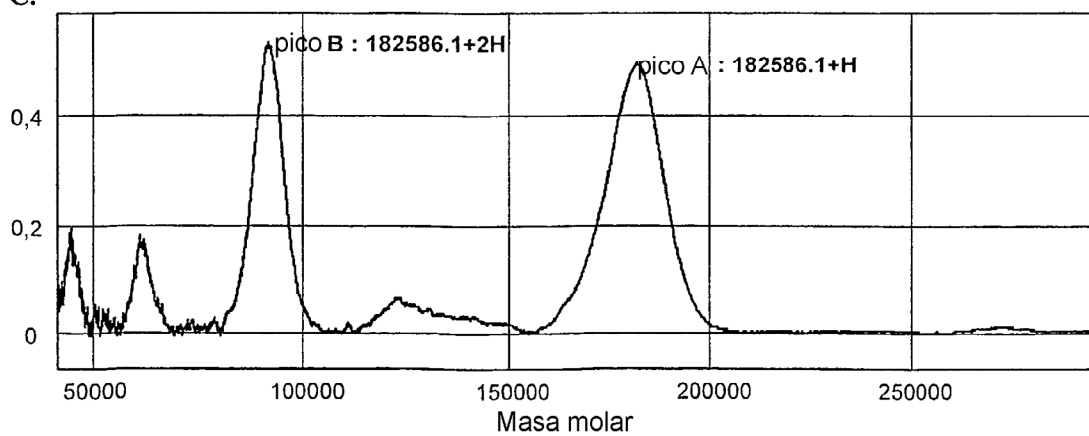


FIGURA 27 A-C

D.

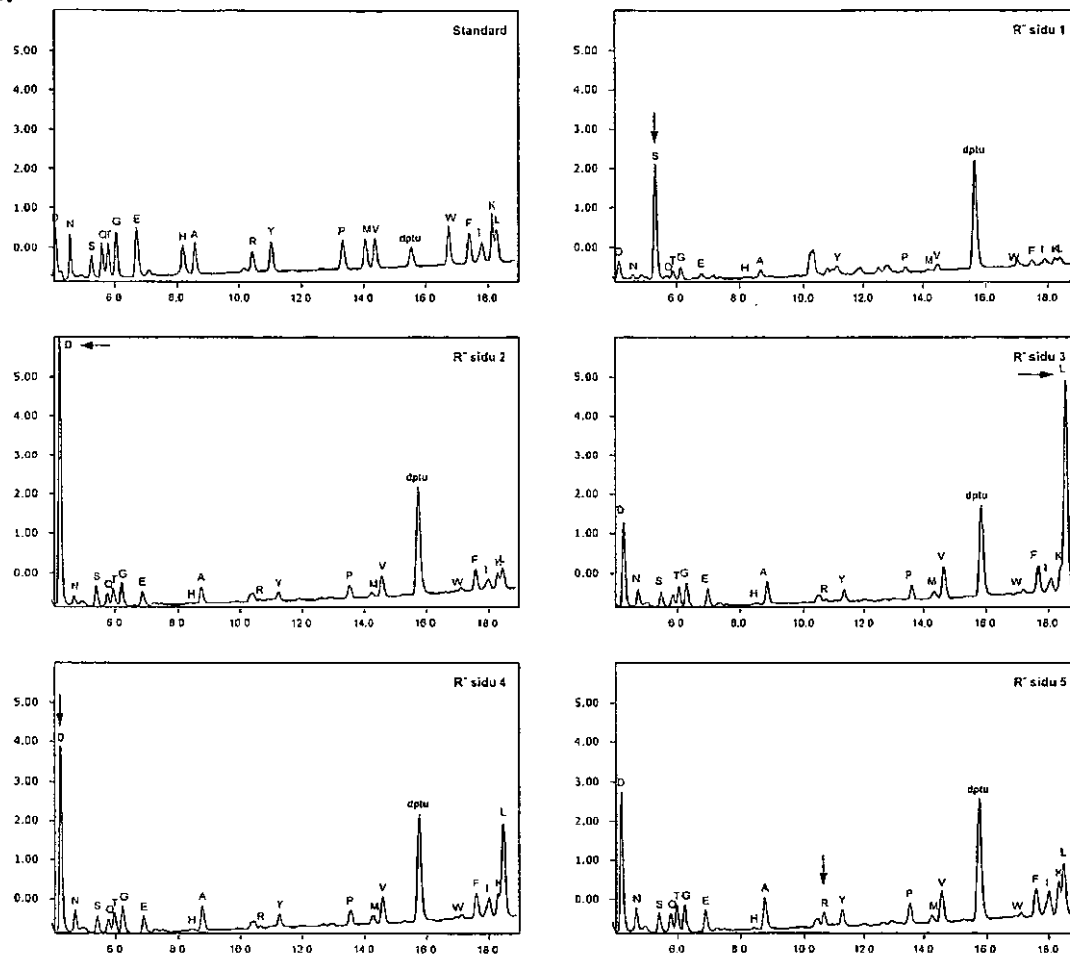
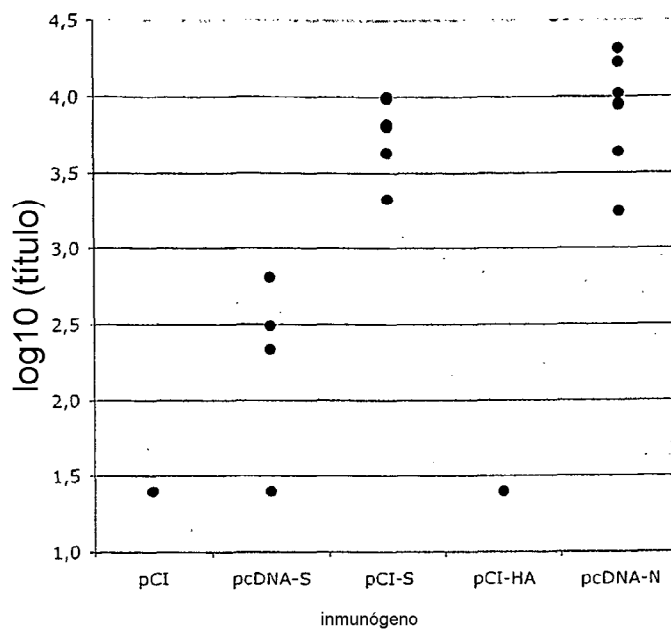


FIGURA 27 D

A.



B.

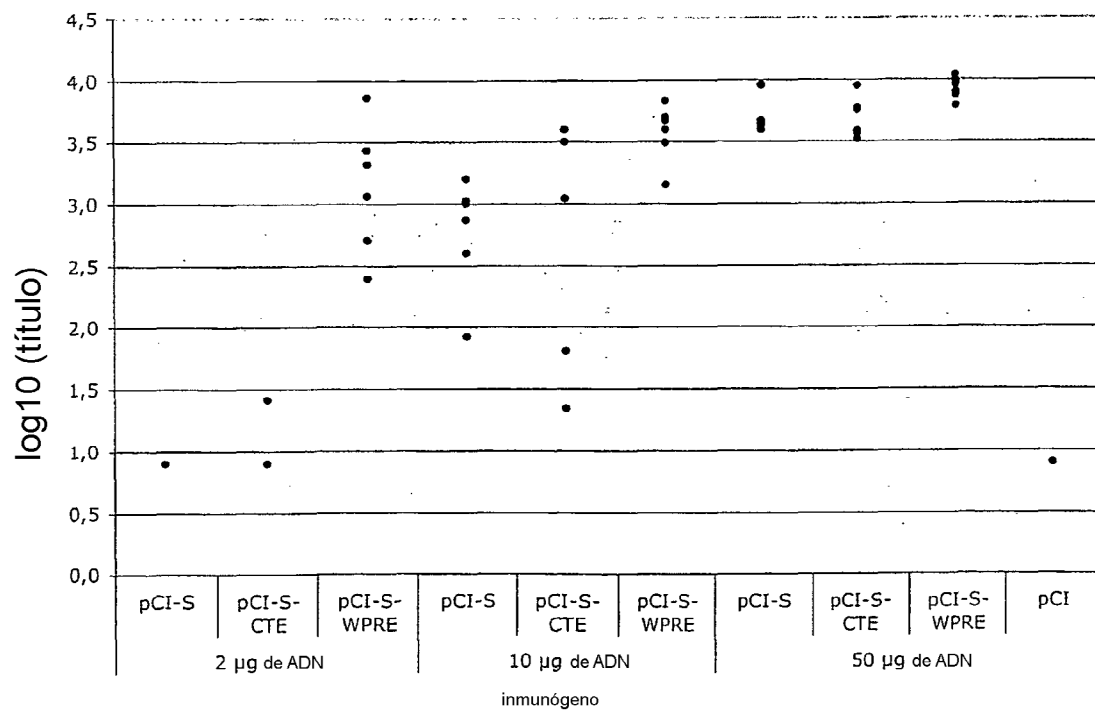
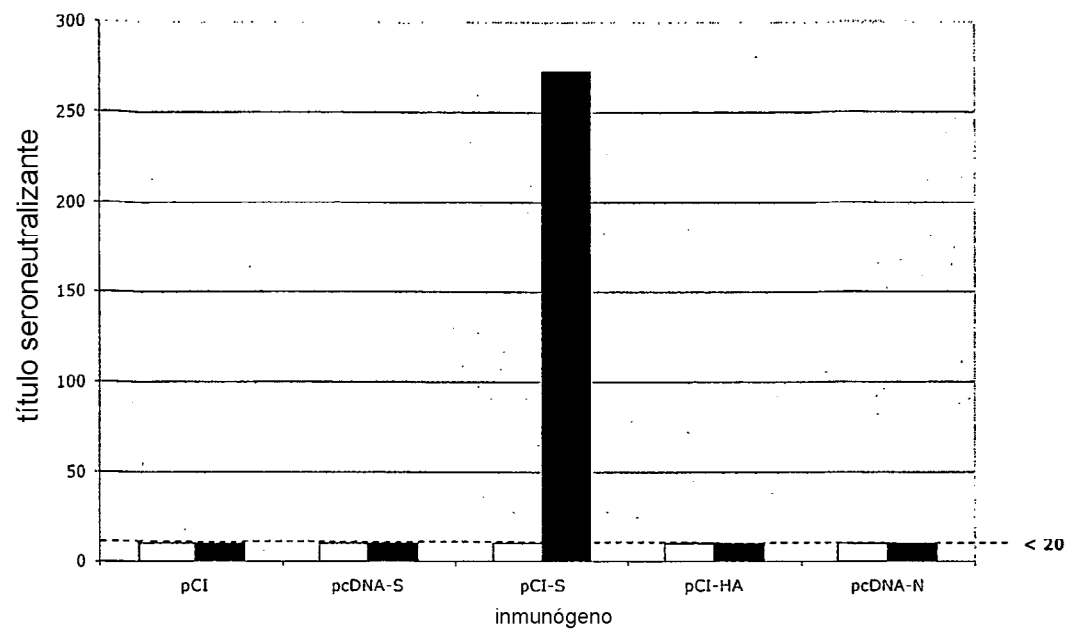


FIGURA 28

A.



B.

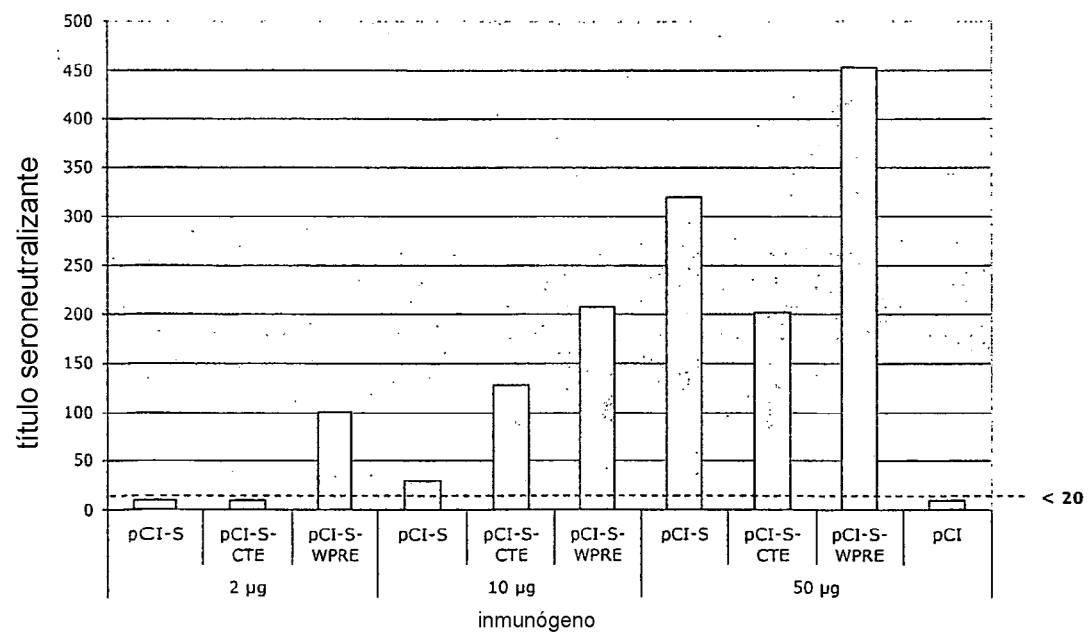


FIGURA 29

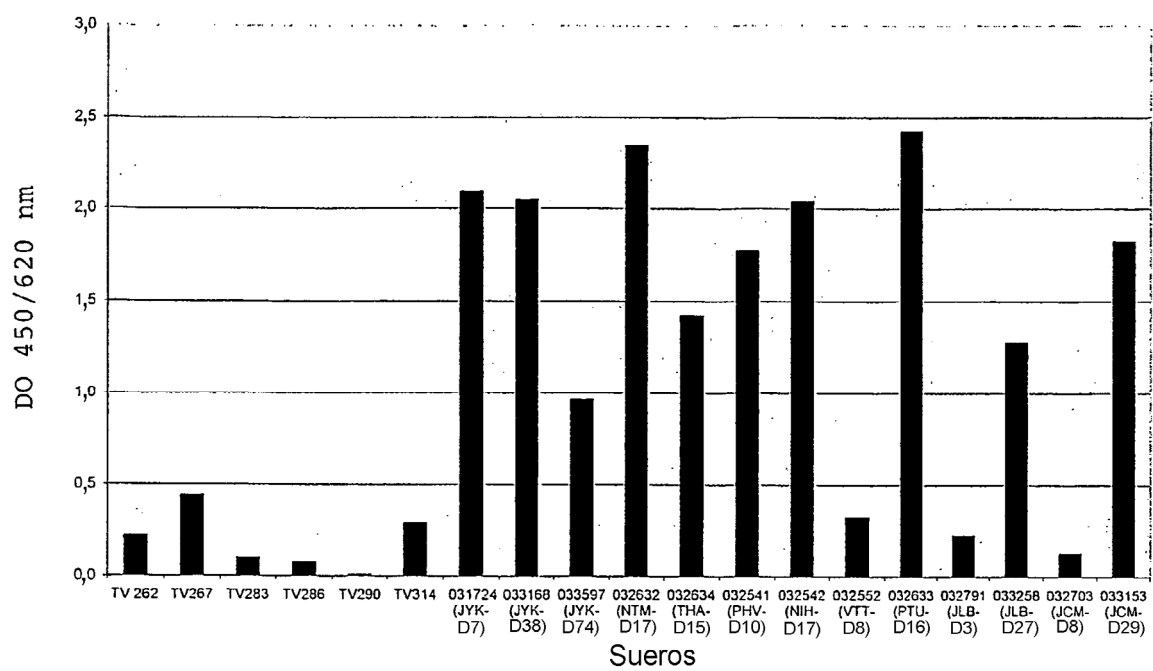


FIGURA 30

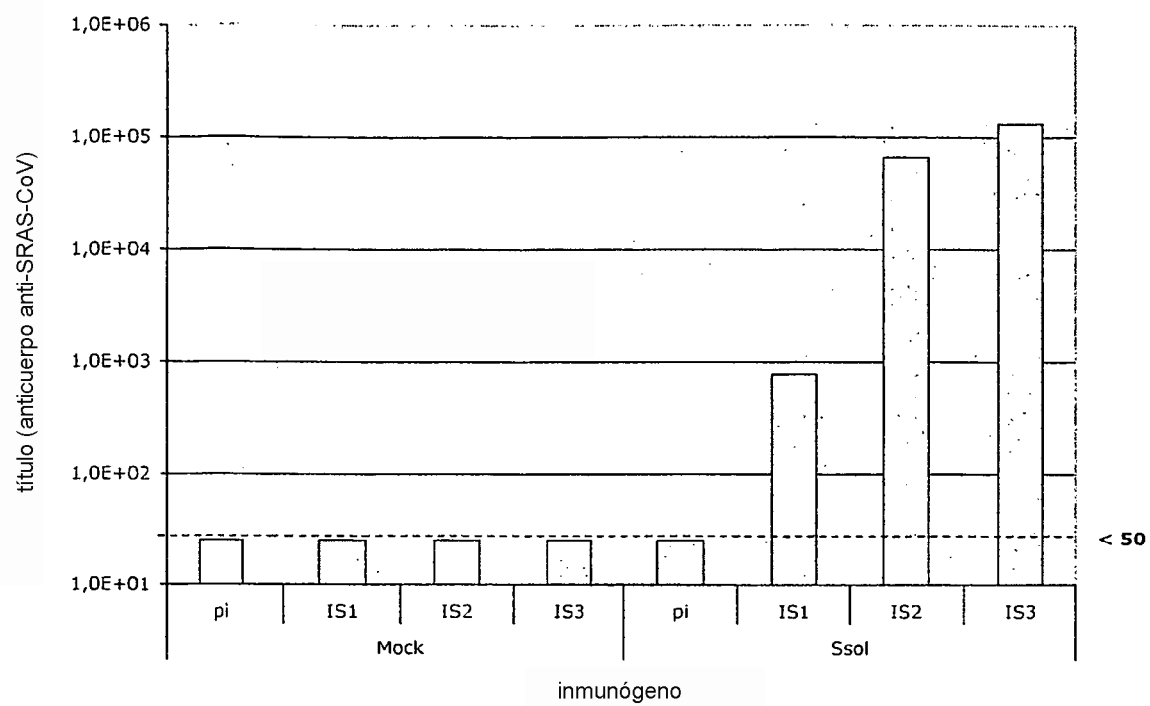


FIGURA 31

I-3059 S-040530	1	CTCTTCTGGAAAAAGGTAGGCTTATCATTAGAGAAAAACAACAGAGTTGTGGTTTCAAGTG
I-3059 S-040530	61 1	ATATTCTTGTTAACAACCTAAACGAACATGTTTATTTTCTTATTATTTCTTACTCTCACTA GG"T"C"C""""C"C"C""C"GC"G"C""G"C""G"C""C"
I-3059 S-040530	121 44	GTGGTACTGACCTTGACCGGTGCACCACTTTTGATGATGTTCAAGCTCCTAATTACACTC "C"C"C""""G""""""""""C"C"C"C""C"G"G"C""C"C""""C"
I-3059 S-040530	181 104	AACATACTT_CATCTATGAGGGGGGTTTACTATCCTGATGAAATTTTAGATCAGACACT "G"C"C"CAG"G"_"_""C""C""G""""C"C"C""G"C""C"GAGC""""C"
I-3059 S-040530	240 163	CTTTATTTAACTCAGGATTTATTTCTTCCATTTTATTTCTAATGTTACAGGGTTTCATACT ""G""CC"G"C""""CC"G"C""G"C""C""CAGC""C""G""C"C""C"C""C"
I-3059 S-040530	300 223	ATTAATCATACGTTTGGCAACCCGTGCATACCTTTTAAGGATGGTATTTATTTTGCTGCC ""C"C"C""C"C""C""""C""G"C"C""C""C""C""C""C"C""C""C""C""C"
I-3059 S-040530	360 283	ACAGAGAAATCAAATGTTGTCCGTGGTTGGGTTTGTGTTCTACCATGAACAACAAGTCA ""C""""GAGC""C""G""G""G""C""""G""C""CAGC""""""""AGC
I-3059 S-040530	420 343	CAGTCGGTGATTATTATTAAACAATTCTACTAATGTTGTTATACGAGCATGTAACCTTTGAA ""AGC""""C"C"C""C""""CAGC""C"C""G""G""C""G""C""C""C""C""G
I-3059 S-040530	480 403	TTGTGTGACAACCCCTTTCTTTGCTGTTTCTAAACCCATGGGTACACAGACACATACTATG C""C"C""""C""""C""C""G""C""""C""C""C""C""C""C""C""C""C""C"
I-3059 S-040530	540 463	ATATTCGATAATGCATTTAATTGCACTTTGAGTACATATCTGATGCCTTTTCGCTTGAT ""C""""C"C"C""C""C""C""C""C""C""C""C""C""C""CAGC""C""C""CAGC""G"C"
I-3059 S-040530	600 523	GTTTCAGAAAAGTCAGGTAATTTTAAACACTTACGAGAGTTTGTGTTTAAAAATAAGAT ""GAGC""G""AGC""C"C"C""G""C""G""G""C""C""C""C""C""G""C""G""C"
I-3059 S-040530	660 583	GGGTTTCTCTATGTTTATAAGGGCTATCAACCTATAGATGTAGTTGCTGATCTACCTTCT ""C"C""G""C""G""C""C""C""C""C""C""C""G""C""C""C""G""GA""A""C""G""CAGC
I-3059 S-040530	720 643	GGTTTAAACACTTTGAAACCTATTTTAAAGTTGCCCTCTGGTATTAACATTACAAATTTT ""C"C""""CC""""G""C""C""C""C""C""C""C""C""C""G""C""C""C""C""C""C"
I-3059 S-040530	780 703	AGAGCCATTCTTACAGCCTTTTCACCTGCTCAAGACATTTGGGGCACGTGAGCTGCAGCC C"G""""C""G""C""C""C""AGC""C""C""G""C""C""C""C""C""CAGC""C""C""C"
I-3059 S-040530	840 763	TATTTGTTGGCTATTTAAAGCCAACCTACATTTATGCTCAAGTATGATGAAAATGGTACA ""C"C""G""C""C""C""C""G""C""C""C""C""C""C""C""C""C""G""C""C""C""C"
I-3059 S-040530	900 823	ATCACAGATGCTGTTGATTGTTCTCAAAATCCACTTGCTGAACCTCAATGCTCTGTTAAG ""C""C""C""C""G""C""CAGC""G""C""C""G""C""G""G""G""AGC""G""C"
I-3059 S-040530	960 883	AGCTTTGAGATTGACAAAGGAATTTACCAGACCTCTAATTTCAGGGTTGTTCCCTCAGGA ""C""C""C""C""C""C""G""C""C""C""C""C""C""C""C""C""C""C""C""C""C"
I-3059 S-040530	1020 943	GATGTTGTGAGATTCCTAATATTACAACTTGTGTCCTTTTGGAGAGGTTTAAATGCT ""C"
I-3059 S-040530	1080 1003	ACTAAATCCCTTCTGTCTATGCATGGGAGAGAAAAAAATTTCTAATTTGTGTTGCTGAT ""C""G""C""CAGC""G""C""C""C""C""C""C""G""G""G""CAGC""C""C""G""C""C"
I-3059 S-040530	1140 1063	TACTCTGTGCTCTACAACCTCAACATTTTTTCAACCTTTAAGTGCTATGGCGTTTCTGCC ""AGC""""C""C""C""C""C""C""C""CAGC""C""C""C""C""C""C""C""C""C"
I-3059 S-040530	1200 1123	ACTAAGTTGAATGATCTTTGCTTCTCCAATGCTATGCAGATTCTTTGTAGTCAAGGGA ""C""C""C""C""C""C""G""C""C""AG""C""G""C""C""C""CAGC""C""G""G""C""C"
I-3059 S-040530	1260 1183	GATGATGTAAGACAAATAGCGCCAGGACAACTGGTGTATTGCTGATTATAATTATAAA ""C""C""G""C""C""C""C""C""C""C""C""C""C""G""C""C""C""C""C""C""C""C"

FIGURA 32.1

ES 2 396 127 T3

[illegible]

FIGURA 32.2

I-3059	2697	GGTGCTGGCGCTGCTCTTCAAATACCTTTTGCTATGCAATGGCATATAGSTTCAATGGC
S-040530	2620	" "A""C""A"" "C""C""G""G""C""C""C""C"" ""G"" ""C""CC"" ""C"" "
I-3059	2757	ATTGGAGTTACCCAAATGTTCTCTATGAGAACC AAAACAATCGCCAACCAATTTAAC
S-040530	2680	" "C""C""G"" " " "G""C""G""G""C"" "" ""G""G""G"" "" ""G""C"" "
I-3059	2817	AAGGCGATTAGTCAAATTCAGAATCACTTACAACAACATCAACTGCATTGGGCAAGCTG
S-040530	2740	" "" ""C""C"" " "C""G""C""G""GAGC""G""C"" "" ""CAGC""C""CC"" "" "" "" "
I-3059	2877	CAAGACGTGTGTTAACCAGAATGCTCAAGCATTAAACACACTTGTTAACAACCTTAGCTCT
S-040530	2800	" "G"" "" ""G"" "G"" "" "" ""C""C""G""CC"G"" "" ""C""G""G""G""G"" "AGC
I-3059	2937	AATTTTGGTGCAATTTCAAGTGTGCTAAATGATATCCTTTTCGCGACTTGATAAAGTCGAG
S-040530	2860	" "C""C""C"" "C""CAGCTC"" "" ""G""C""C"" "" ""GAGCA"G""G""C"" "" ""G"" "
I-3059	2997	GCGGAGGTACAAATTGACAGGCTAATTACAGGCAGACTTCAAAGCCTTCAAACCTATGTA
S-040530	2920	" "C""A""G"" "G""C"" "" ""G""C""C""AC"C""G""GTC"" "G""G"" "" ""C""G"
I-3059	3057	ACACAACAAC TAATCAGGGCTGCTGAAATCAGGGCTTCTGCTAATCTTGCTGCTACTAAA
S-040530	2980	" "C""G""G"" "G"" "" ""A""C""C""G"" ""C"" ""CAGC""C"" "" ""G""C""C""C""G"
I-3059	3117	ATGTCTGAGTGTTCTTGACAATCAAAAAGAGTTGACTTTTGTGGAAAGGGCTACCAC
S-040530	3040	" ""AGC"" "" ""C""G""G""C""GAGC""G"" "" ""G"" "" ""C""C""C"" "" ""T"" "
I-3059	3177	CTTATGTCCTTCCCACAAGCAGCCCCGCATGGTGTGTCTTCTACATGTCACGTATGTG
S-040530	3100	" "G"" "AG"" "" ""C""G""C"" "" ""C""C""C""G""G"" "" ""G""C""G""C""C"" "
I-3059	3237	CCATCCAGGAGAGGAAC TTACCACAGCGCCAGCAATTTGTCATGAAGGCCAAAGCATAC
S-040530	3160	" "TAG"" "" "" "C"" "" "" "" ""C""C"" "" ""C""C""C""G"" "" ""G""C"" "
I-3059	3297	TTCCTTCGTGAAGGTGTTTTTGTGTTTAATGGCACTTCTTGSTTTATTACACAGAGGAAC
S-040530	3220	" "" ""C""G"" "G""C""G""C"" "" "" ""C"" "" ""CAGC"" "" ""C""C""C""C"" "
I-3059	3357	TTCTTTTCTCCACAAATAATTACTACAGACAATACATTTGTCTCAGGAAATGTGATGTC
S-040530	3280	" "" ""CAGC""C""G""C""C""C"" "" "" "" ""C""C""C""G""C""C""C"" "" "" ""G"
I-3059	3417	GTTATTGGCATCATTAACAACACAGTTTATGATCCTCTGCAACCTGAGCTTGACTCATTC
S-040530	3340	" "G""C"" "" "" ""C""T"" "" ""C""G""C""C""C"" "" ""G""C"" "" ""G"" "AGC"" "
I-3059	3477	AAAGAAGAGCTGGACAAGTACTTCAAAAATCATACATCACCAGATGTTGATCTTGGCGAC
S-040530	3400	" "G""G"" "" "" "" ""A"" "" "" ""G""C""C""C""C""C""C""G""C""G"" "" ""T"
I-3059	3537	ATTTCAGGCATTAAACGCTTCTGTGCTCAACATTCAAAAAGAAATTGACCGCCTCAATGAG
S-040530	3460	" "CAGC"" "" ""C"" "" ""C""C""G""G"" "" ""C""G""G""G""C"" "" ""A""A""G""C""A"
I-3059	3597	GTCGCTAAAAATTTAAATGAATCACTCAT TGACCTTCAAGAAT TGGGAAAATATGAGCAA
S-040530	3520	" "G""C""G"" "CC"G""C""GAGC""G""C"" "" ""G""G""GC"" "" ""C""G""C"" "" ""G"
I-3059	3657	TATATTAAATGGCCTTGGTATGTTTGGCTCGGCTTCATTGCTGGACTAATTGCCATCGTC
S-040530	3580	" "C""C""G"" " " "C"" "" ""C""G"" "" ""G"" "" "" ""C""C""C""G""C"" "" "" ""G"
I-3059	3717	ATGTTTACAATCTTGCTTTGTGTCATGACTAGTTGTTGCACTTGCCCTCAAGGGTGCATGC
S-040530	3640	" "" ""G""C"" " "C"" "" ""G""C"" "" "" ""C""C""C""T""C"" "" ""G""A""C""C"" "
I-3059	3777	TCTTGTTGGTTCTTGCTGCAAGTTTGATGAGGATGACTCTGAGCCAGTTCTCAAGGGTGTG
S-040530	3700	AGC"" "" ""CAGC"" "" "" "" ""C""C"" "" ""C"" "AGC"" "" ""C""G""G"" "" ""C""G"
I-3059	3837	AAATTACATTACACATAAACGAACCTTATGGATTTGTTTATGAGATTTTTTACTCTTGGAT
S-040530	3760	" "GC"G""C"" " " "C""G""T _ "" ""CGA"" _
I-3059	3897	CAATTACTGCACAGCCAGTAAAAATTGACAATGCTTCTCCTGCAAGT
S-040530		

FIGURA 32.3

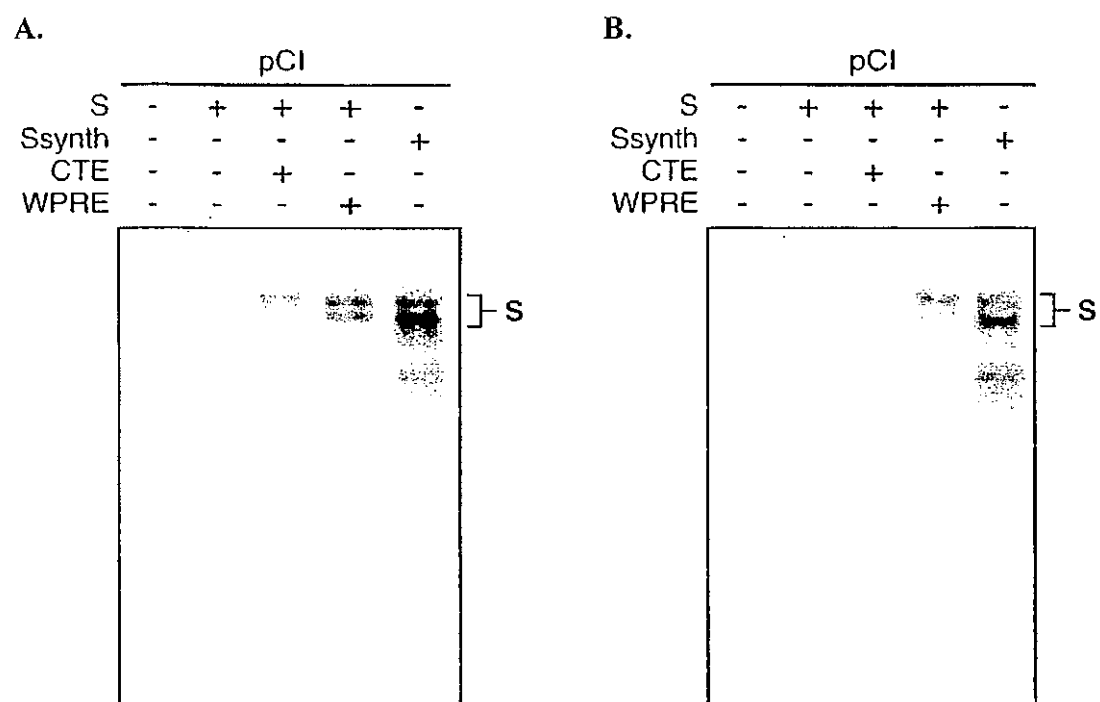
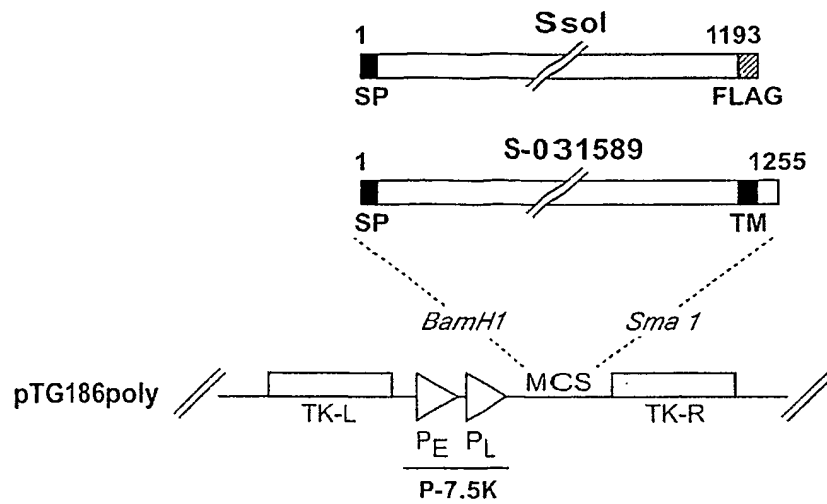
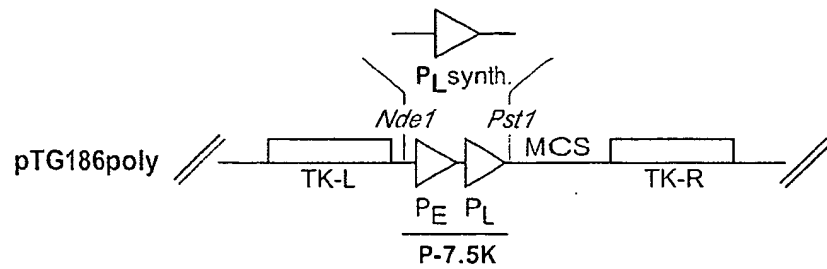


FIGURA 33

A.



B.



C.

$\frac{\text{CATATG}}{\text{NdeI}}$ AGC $\frac{[\text{T}]_{20}\text{GGCATATAAATA}}{\text{promotor 480}}$ GACTC $\frac{\text{GGCGCGCC}}{\text{AscI}}$ AT $\frac{\text{CTGCAG}}{\text{PstI}}$

FIGURA 34 A-C

D.

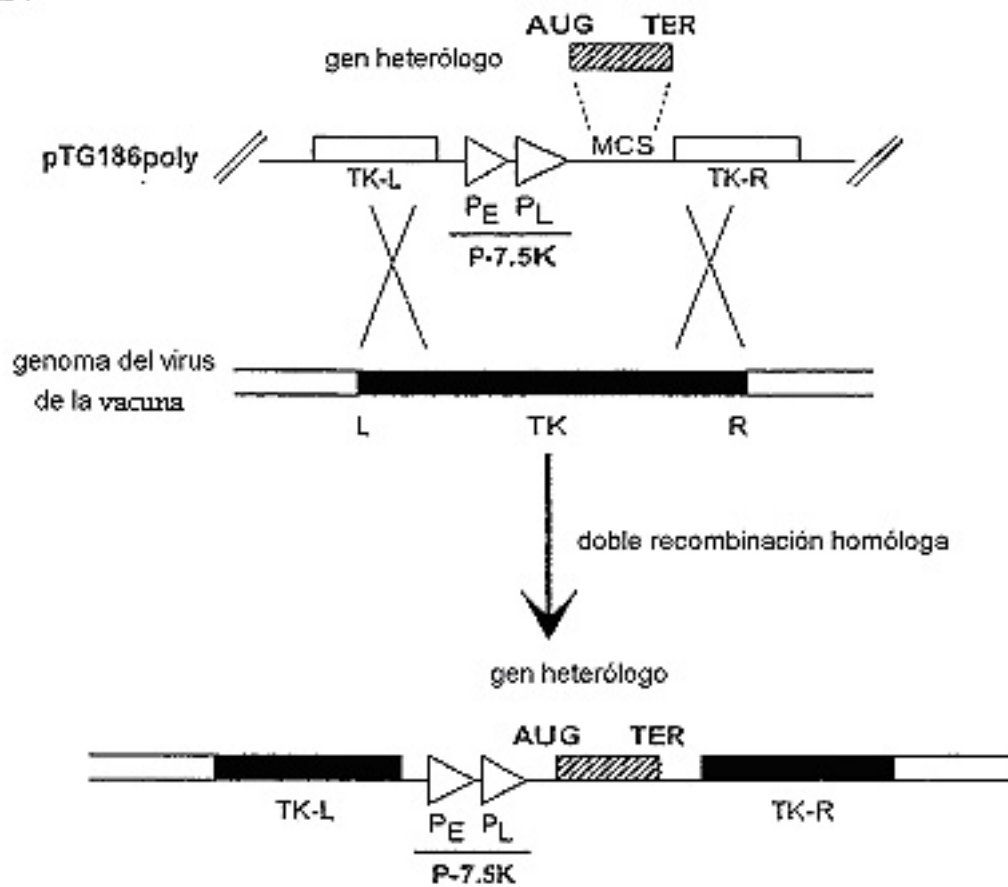
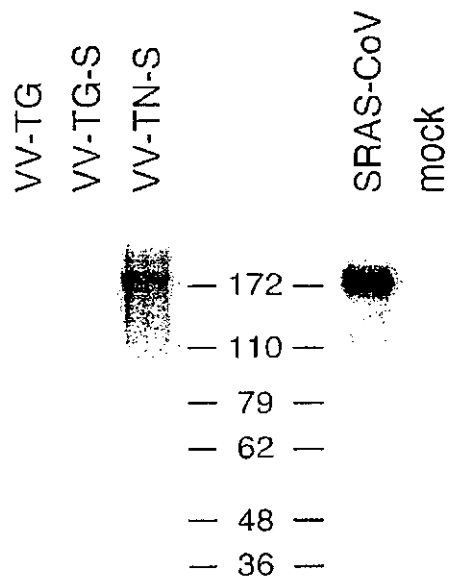


FIGURA 34 D

A.



B.

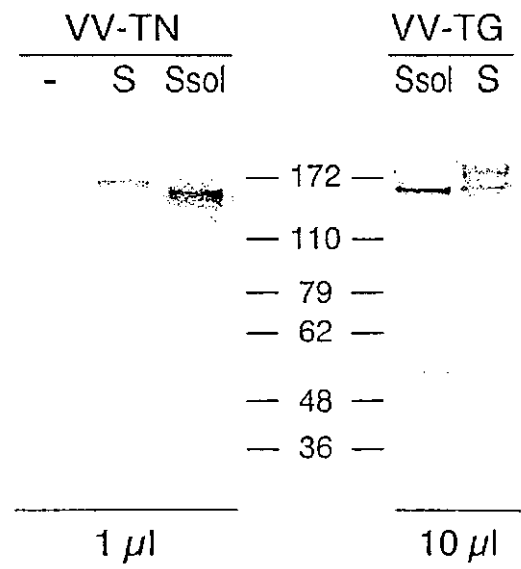
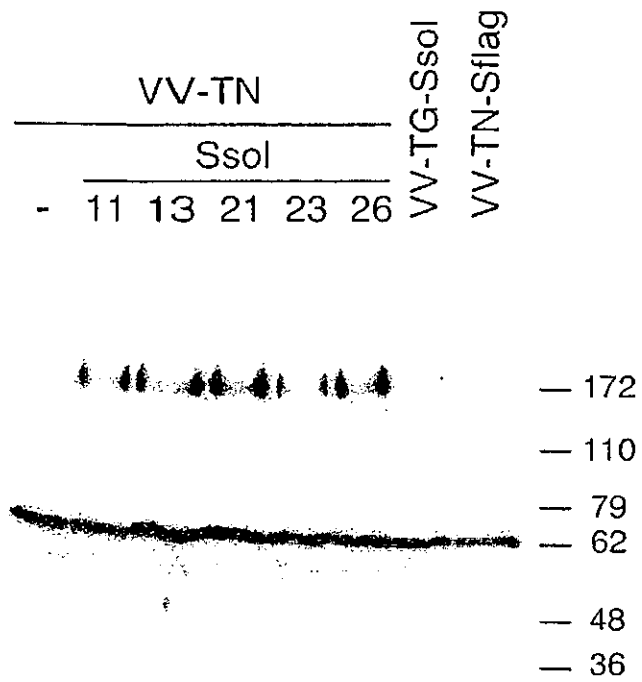


FIGURA 35

A.



B.

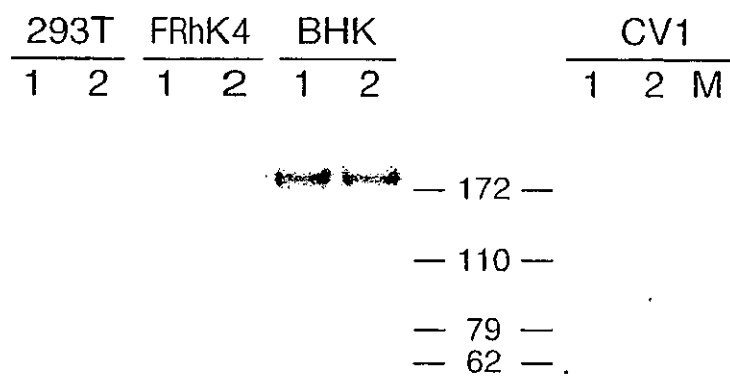


FIGURA 36

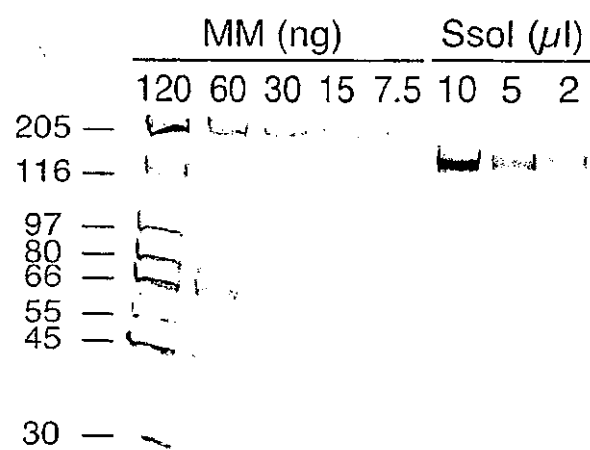


FIGURA 37

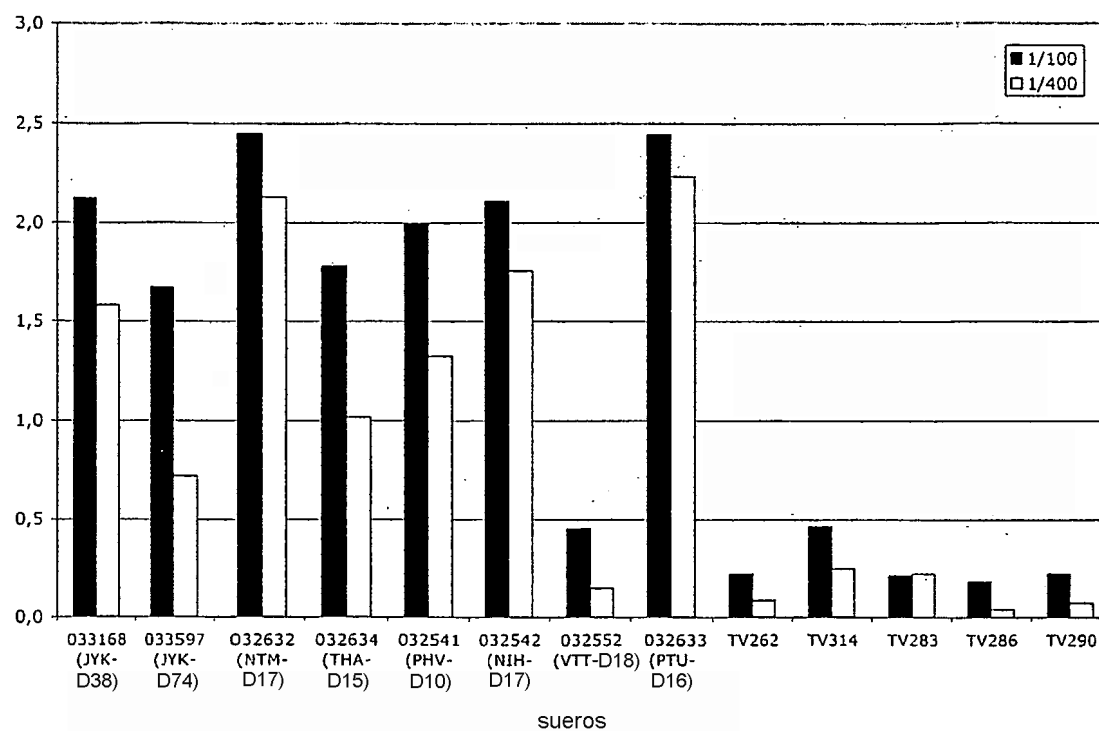
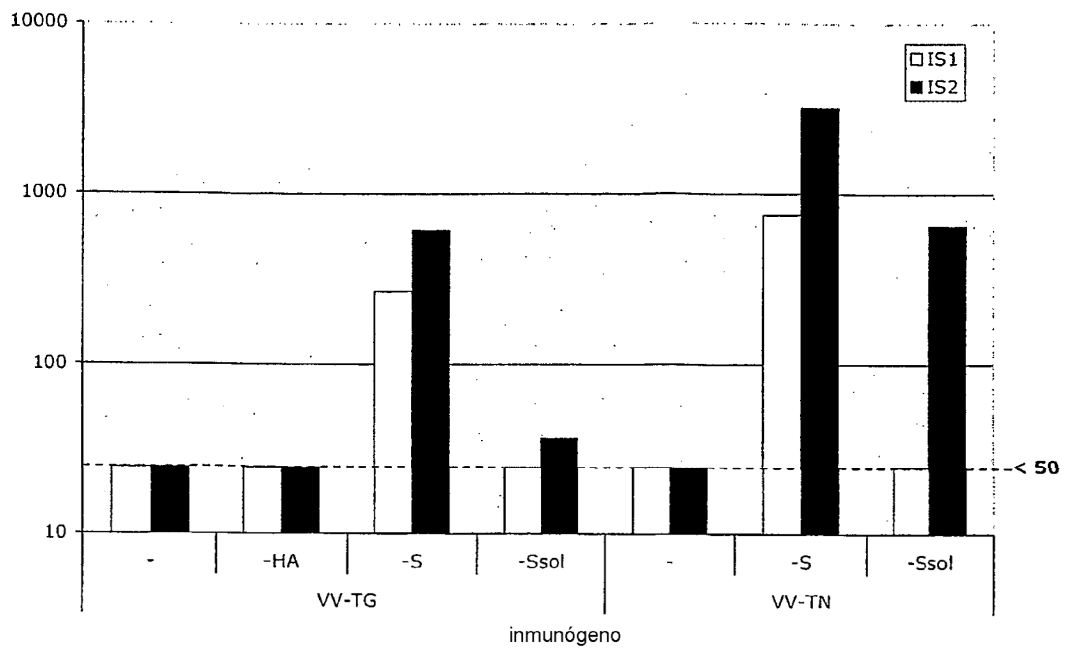


FIGURA 38

A.



B.

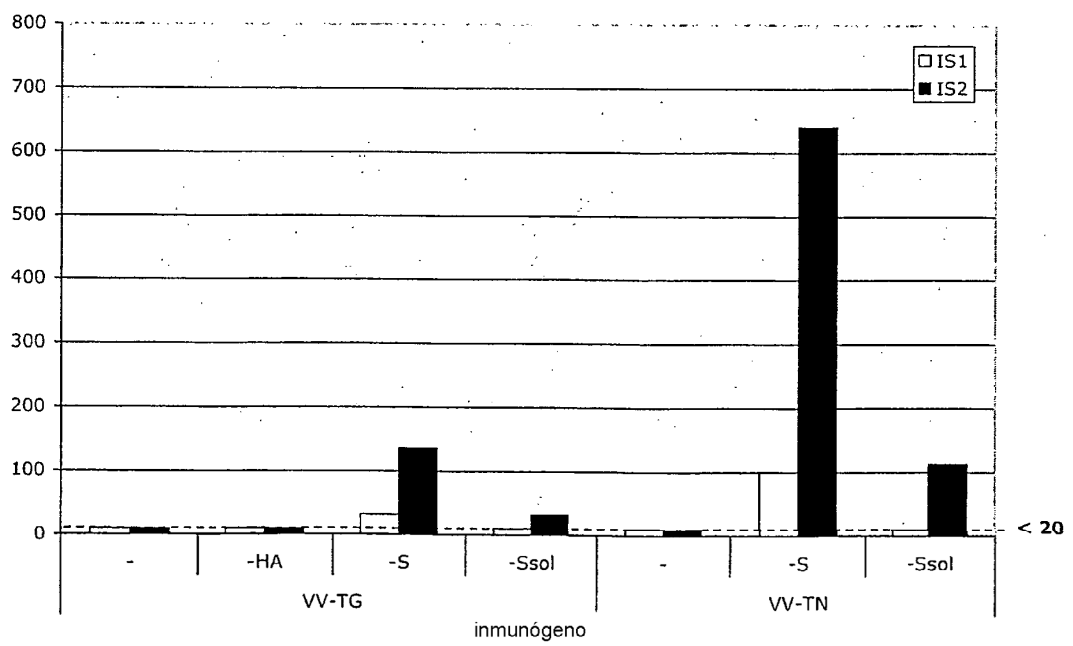


FIGURA 39

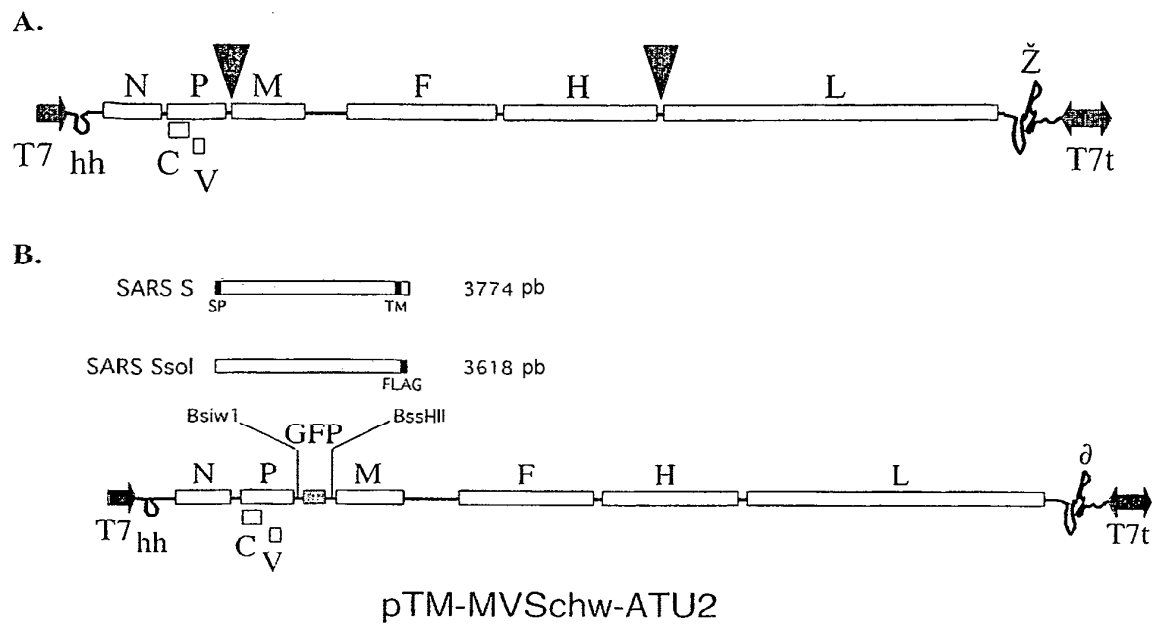


FIGURA 40

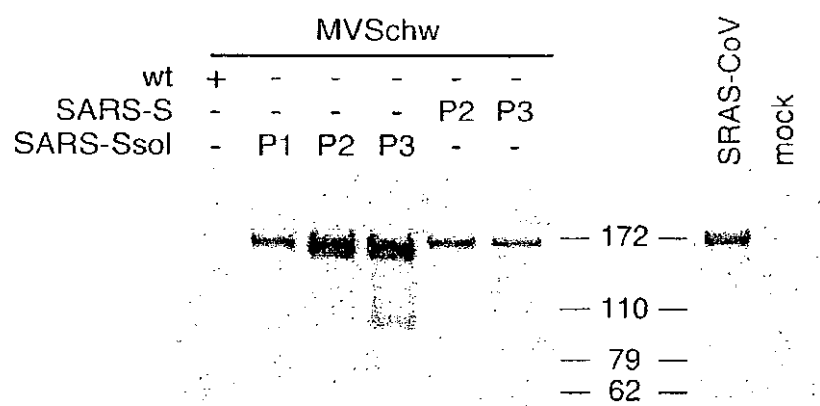


FIGURA 41

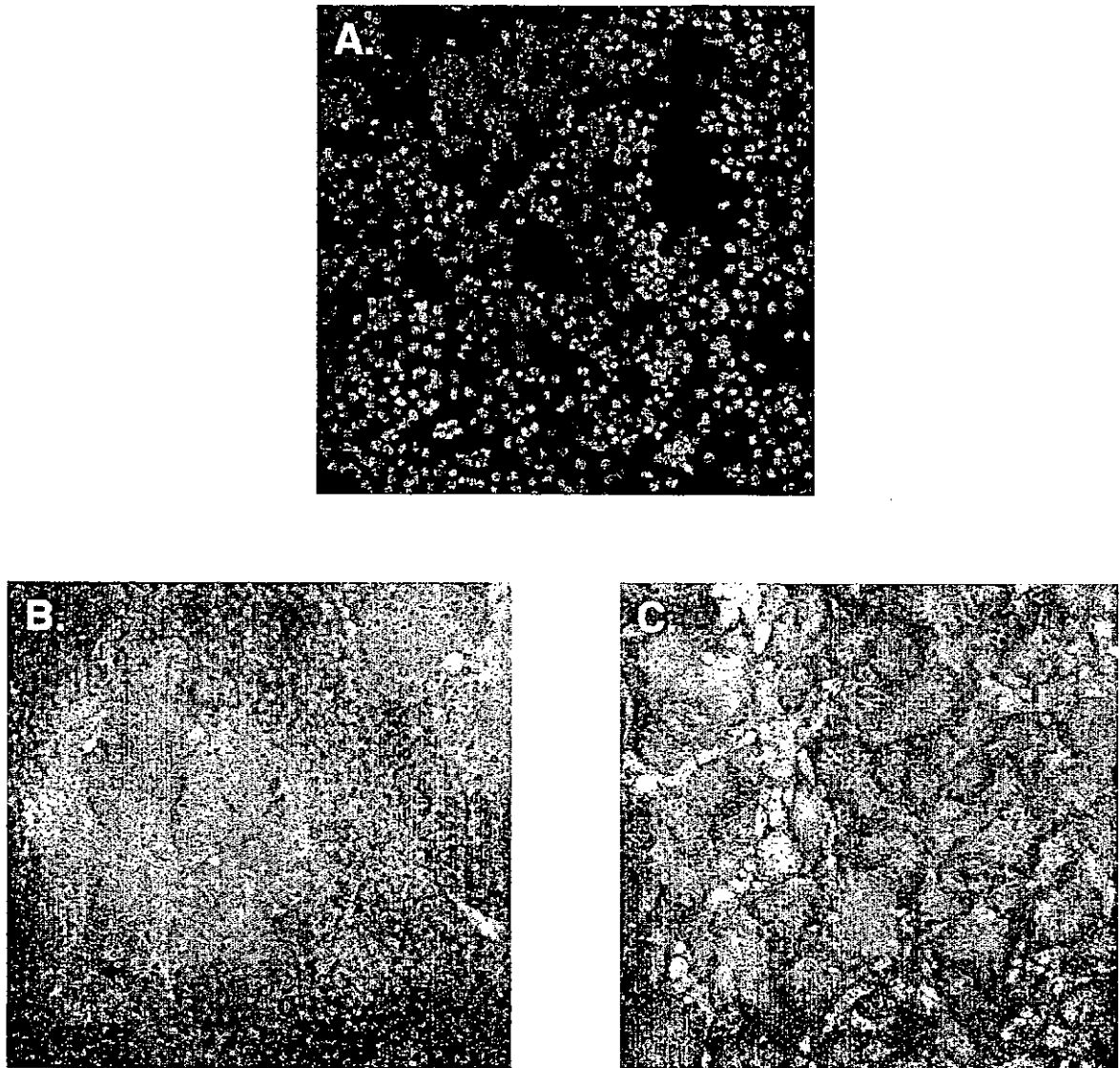


FIGURA 42

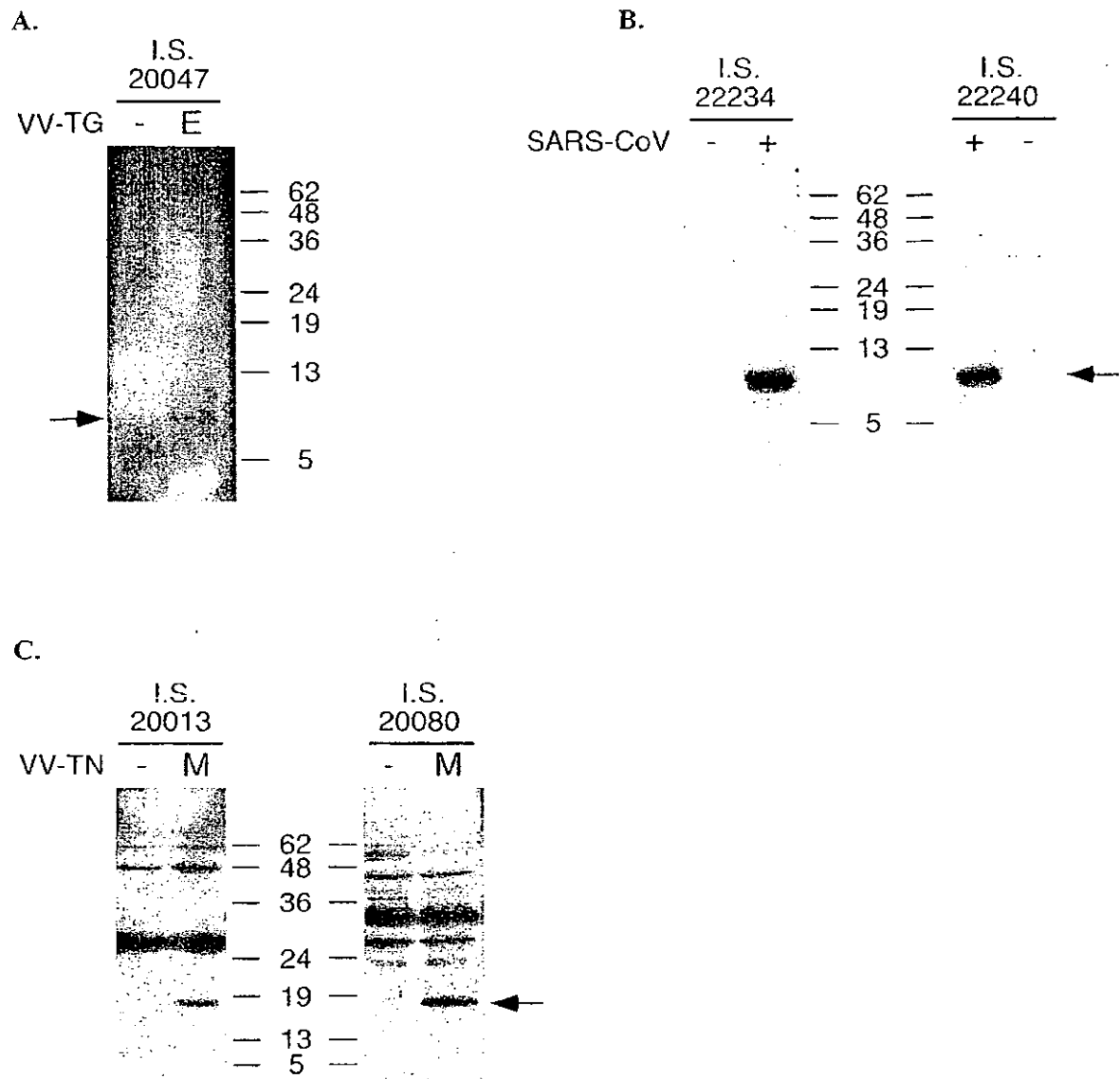


FIGURA 43