



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 331 251**

51 Int. Cl.:
C12N 15/48 (2006.01)
C12Q 1/70 (2006.01)
C07K 14/15 (2006.01)
A61K 31/70 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **05017735 .1**
96 Fecha de presentación : **07.07.1998**
97 Número de publicación de la solicitud: **1612270**
97 Fecha de publicación de la solicitud: **04.01.2006**

54 Título: **Material nucleico retrovítico y fragmentos nucleotídicos, asociados a la esclerosis en placas, con fines de diagnóstico, profilácticos y terapéuticos.**

30 Prioridad: **07.07.1997 FR 97 08816**

45 Fecha de publicación de la mención BOPI:
28.12.2009

45 Fecha de la publicación del folleto de la patente:
28.12.2009

73 Titular/es: **BIOMERIEUX**
Chemin de l'Orme
69280 Marcy-L'Etoile, FR

72 Inventor/es: **Parahnos-Baccala, Glaucia;**
Komurian-Pradel, Florence;
Bedin, Frédéric;
Sodoyer, Mireille;
Ott, Catherine;
Perron, Hervé;
Mandrand, Bernard y
Mallet, Francois

74 Agente: **Curell Suñol, Marcelino**

ES 2 331 251 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

Material nucleico retrovívico y fragmentos nucleotídicos, asociados a la esclerosis en placas, con fines de diagnóstico, profilácticos y terapéuticos.

La esclerosis en placas (SEP) es una enfermedad desmielinizante del sistema nervioso central (SNC) cuya causa completa sigue siendo desconocida.

Numerosos estudios han apoyado la hipótesis de una etiología vírica de la enfermedad, pero ninguno de los virus conocidos ensayados ha resultado ser el agente causante buscado: un repaso de los virus buscados desde hace varios años en la SEP ha sido realizado por E. Norrby y R.T. Johnson.

Recientemente, un retrovirus, diferente de los retrovirus humanos conocidos, se ha aislado en unos pacientes que padecen SEP. Los autores han podido demostrar asimismo que este retrovirus podía transmitirse *in vitro*, que unos pacientes que padecen SEP producían unos anticuerpos susceptibles de reconocer unas proteínas asociadas a la infección de las células leptomeníngeas por este retrovirus, y que la expresión de este último podría ser estimulada en gran medida por los genes inmediatos-precoces de ciertos virus del herpes.

Todos estos resultados justifican el papel en la SEP de por lo menos un retrovirus desconocido o de un virus que tiene una actividad transcriptasa inversa (RT) detectable según el método publicado por H. Perron y calificado de actividad "RT de tipo LM7".

Los estudios del solicitante han permitido obtener dos líneas continuas de células infectadas por unos aislados naturales que proceden de dos pacientes diferentes que padecen SEP, mediante un procedimiento de cultivo tal como se describe en el documento WO-A-93 20188. Estas dos líneas derivadas de células de Plexus-coroideas humanas, denominadas LM7PC y PLI-2 se han depositado en la E.C.A.C.C. respectivamente el 22 de julio de 1992 y el 8 de enero de 1993, con los números 92 072201 y 93 010817, de acuerdo con las disposiciones del Tratado de Budapest. Por otro lado, los aislados víricos que poseen una actividad RT de tipo LM7 también se han depositado en la E.C.A.C.C. con la denominación global de "cepas". La "cepa" o aislado hospedado por la línea PLI-2, denominada POL-2, se ha depositado en la E.C.A.C.C. el 22 de julio de 1992 con el nº V92072202. La "cepa" o aislado hospedado por la línea LM7PC, denominada MS7PG, se ha depositado en la E.C.A.C.C. el 8 de enero de 1993 con el nº V93010816.

A partir de los cultivos y de las muestras citados anteriormente, caracterizados por unos criterios biológicos y morfológicos, se ha procedido después a caracterizar el material nucleico asociado a las partículas víricas producidas en estos cultivos.

Las porciones de genoma ya caracterizadas se han usado para realizar unos ensayos de detección molecular del genoma vírico y unos ensayos inmunoserológicos, utilizando las secuencias de aminoácidos codificadas por las secuencias nucleotídicas del genoma vírico, para detectar la respuesta inmunitaria dirigida contra unos epítomos asociados a la infección y/o a la expresión vírica.

Estas herramientas han permitido ya confirmar una asociación entre la SEP y la expresión de las secuencias identificadas en las patentes citadas más adelante. Sin embargo, el sistema vírico descubierto por el solicitante, se asemeja a un sistema retrovívico complejo. En efecto, las secuencias encontradas encapsidadas en las partículas víricas extracelulares producidas por los diferentes cultivos de células de pacientes que padecen SEP, muestran claramente que existe una co-encapsidación de genomas retrovívicos parecidos, pero diferentes del genoma retrovívico "salvaje" que produce las partículas víricas infectantes. Este fenómeno se ha observado entre unos retrovirus replicativos y unos retrovirus endógenos, la noción de retrovirus endógeno es muy importante en el contexto del presente descubrimiento porque, en el caso de MSR-1, se ha observado que unas secuencias retrovívicas endógenas que comprenden unas secuencias homólogas al genoma MSR-1 existen en el ADN humano normal. La existencia de elementos retrovívicos endógenos (ERV) parecidos a MSR-1 mediante la totalidad o parte de su genoma, explica el hecho de que la expresión del retrovirus MSR-1 en las células humanas pueda interactuar con unas secuencias endógenas cercanas. Estas interacciones se vuelven a encontrar en el caso de retrovirus endógenos patógenos y/o infecciosos (por ejemplo ciertas cepas ecótropas del Murine Leukaemia virus), en el caso de retrovirus exógenos cuya secuencia nucleotídica se puede encontrar parcial o totalmente, en forma de ERV, en el genoma del animal hospedante (por ejemplo, virus exógeno del tumor mamario del ratón transmitido por la leche). Estas interacciones consisten principalmente en (i) una transactivación o co-activación de ERVs por el retrovirus replicativo, (ii) una encapsidación "ilegítima" de ARN parecidos a ERVs, o ERVs -incluso de ARN celulares- que poseen simplemente unas secuencias de encapsidación compatibles, en las partículas retrovívicas producidas por la expresión de la cepa replicativa, a veces transmisibles y a veces con una patogenicidad propia, y (iii) unas recombinaciones más o menos importantes entre los genomas co-encapsidados, en particular en las fases de transcripción inversa, que conducen a la formación de genomas híbridos, a veces transmisibles y a veces con una patogenicidad propia.

Así, (i) diferentes secuencias parecidas a MSR-1 se han encontrado en las partículas víricas purificadas; (ii) el análisis molecular de las diferentes regiones del genoma retrovívico MSR-1 se debe realizar analizando sistemáticamente las secuencias co-encapsidadas, interferentes y/o recombinadas que se generan mediante la infección y/o la expresión de MSR-1, además, ciertos clones pueden tener unas partes de secuencias defectivas producidas por la replicación retrovívica, y los errores de matriz y/o de transcripción de la transcriptasa inversa; (iii) las familias de

secuencias parecidas a una misma región genómica retrovírica son los soportes de una detección diagnóstica global que se puede optimizar por la identificación de regiones invariables entre los clones expresados y mediante la identificación de marcos de lecturas responsables de la producción de polipéptidos antigénicos y/o patógenos que pueden ser producidos sólo por una parte, incluso uno solo, de los clones expresados y en estas condiciones, el análisis sistemático de los clones expresados en una región de un gen determinado permite evaluar la frecuencia de variación y/o de recombinación del genoma MSR-1 en esta región y definir las secuencias óptimas para las aplicaciones, en particular diagnósticas; (iv) la patología provocada por un retrovirus tal como MSR-1 puede ser un efecto directo de su expresión y de las proteínas o péptidos producidos por ello, pero asimismo un efecto de la activación, de la encapsulación, de la recombinación de genomas parecidos o heterólogos y de las proteínas o péptidos producidos por ello; Así, estos genomas asociados a la expresión de y/o a la infección por MSR-1 forman parte integrante de la patogenicidad potencial de este virus y por lo tanto, constituyen unos soportes de detección diagnóstica y unas dianas terapéuticas particulares. Asimismo, cualquier agente asociado, o, co-factor de estas interacciones responsables de la patología en cuestión, tal como MSR-2 o el factor gliotóxico descrito en la solicitud de patente publicada con el n° FR-2 716 198, puede participar en la elaboración de una estrategia global y muy eficaz de diagnóstico, de pronóstico, de seguimiento terapéutico y/o de terapéutica integrada de la SEP en particular, pero también de cualquier otra enfermedad asociada a los mismos agentes.

En este contexto, se ha hecho un descubrimiento paralelo en otra enfermedad autoinmune, la poliartritis reumatoide (PR), que se ha descrito anteriormente en la solicitud de patente francesa registrada con el n° 95 02960. Este descubrimiento muestra que, aplicando unos enfoques metodológicos similares a los usados en los estudios del solicitante sobre la SEP, se ha podido identificar un retrovirus expresado en la PR que comparte las secuencias descritas para MSR-1 en SEP y también, la co-existencia de una secuencia asociada MSR-2 descrita asimismo en la SEP. En lo referente a MSR-1, las secuencias detectadas comúnmente en SEP y PR, se refieren a los genes *pol* y *gag*. En el estado actual de los conocimientos, se pueden asociar las secuencias *gag* y *pol* descritas a las cepas MSR-1 expresadas en estas dos enfermedades.

La presente solicitud de patente tiene por objeto diferentes resultados, suplementarios con relación a los ya protegidos por las solicitudes de patente francesa: n° 92 04322 del 03.04.1992, publicada con el n° 2 689 519;

- n° 92 13447 del 03.11.1992, publicada con el n° 2 689 521;
- n° 92 13443 del 03.11.1992, publicada con el n° 2 689 520;
- n° 94 01529 del 04.02.1994, publicada con el n° 2 715 936;
- n° 94 01531 del 04.02.1994, publicada con el n° 2 715 939;
- n° 94 01530 del 04.02.1994, publicada con el n° 2 715 936;
- n° 94 01532 del 04.02.1994, publicada con el n° 2 715 937;
- n° 94 14322 del 24.11.1994, publicada con el n° 2 727 428;
- n° 94 15810 del 23.12.1994, publicada con el n° 2 728 585; y
- la solicitud de patente WO-97/06260.

La presente invención se refiere en primer lugar a un material nucleico, que puede consistir en un material retrovírico, en el estado aislado o purificado, que puede ser conocido o caracterizado de diferentes maneras:

- comprende una secuencia nucleotídica seleccionada de entre el grupo constituido por (i) SEC ID n°: 130; (ii) la secuencia complementaria a la secuencia (i); y (iii) las secuencias equivalentes a las secuencias (i) o (ii), en particular las secuencias que presentan para cualquier sucesión de 100 monómeros contiguos, por lo menos 50%, preferentemente por lo menos 70% de homología con las secuencias (i) o (ii) respectivamente;
- codifica un polipéptido que presenta, para cualquier sucesión contigua de por lo menos 30 aminoácidos, por lo menos 50%, y preferentemente por lo menos 70% de homología, con una secuencia peptídica seleccionada de entre el grupo constituido por las SEC ID n°: 135 y SEC ID n°: 137;
- el extremo 5' de su gen *pol* empieza en el nucleótido 1419 de SEC ID n°: 130;
- el extremo 3' de su gen *gag* termina en el nucleótido 1418 de SEC ID n°: 130;
- el material nucleico retrovírico tal como se ha definido anteriormente está en particular asociado a por lo menos una enfermedad autoinmune tal como la esclerosis en placas o la poliartritis reumatoide.

ES 2 331 251 T3

La invención se refiere asimismo a un fragmento nucleotídico que responde a por lo menos una de las siguientes definiciones:

- 5 - comprende o consiste en una secuencia nucleotídica seleccionada de entre el grupo constituido por (i) la secuencia SEC ID n°: 130; (ii) la secuencia complementaria de la secuencia (i); y (iii) las secuencias equivalentes a las secuencias (i) o (ii), en particular las secuencias que presentan para cualquier sucesión de 100 monómeros contiguos, por lo menos 50%, y preferentemente por lo menos 70% de homología con las secuencias (i) o (ii) respectivamente;
- 10 - comprende o consiste en una secuencia nucleotídica que codifica para un polipéptido que presenta, para cualquier sucesión contigua de por lo menos 30 aminoácidos, por lo menos 50%, preferentemente por lo menos 70% de homología, con una secuencia peptídica seleccionada de entre el grupo constituido por las SEC ID n°: 135 y SEC ID n°: 137.

15 Otros objetos de la presente invención son los siguientes:

- 20 - una sonda nucleica para la detección de un retrovirus asociado a la esclerosis en placas y/o a la poliartritis reumatoide, susceptible de hibridarse específicamente sobre cualquier fragmento definido anteriormente y que pertenece al genoma de dicho retrovirus; comprende ventajosamente de 10 a 100 nucleótidos, preferentemente de 10 a 30 nucleótidos;
- 25 - un cebador para la amplificación mediante polimerización de un ARN o de un ADN de un retrovirus asociado a la esclerosis en placas y/o a la poliartritis reumatoide, que comprende una secuencia nucleotídica idéntica o equivalente a por lo menos una parte de la secuencia nucleotídica de un fragmento definido anteriormente, en particular una secuencia nucleotídica que presenta para cualquier sucesión de 10 monómeros contiguos, por lo menos 50%, preferentemente por lo menos 70% de homología con por lo menos dicha parte de dicho fragmento; preferentemente la secuencia nucleotídica de un cebador de la invención se selecciona de entre las SEC ID n° 126, SEC ID n° 127, SEC ID n° 128, SEC ID n° 129, SEC ID n° 132 y SEC ID n° 133;
- 30 - un ARN o ADN, y en particular un vector de replicación y/o de expresión que comprende un fragmento genómico del material nucleico o un fragmento definido anteriormente;
- 35 - un péptido codificado por cualquier marco de lectura abierto que pertenece a un fragmento nucleotídico definido anteriormente, en particular un polipéptido, por ejemplo oligopéptido que forma o que comprende un determinante antigénico reconocido por unos sueros de pacientes infectados por el virus MSRV-1, y/o en los que el virus MSRV-1 ha sido reactivado; un péptido preferido comprende una secuencia idéntica, parcial o totalmente, o equivalente a una secuencia seleccionada de entre las SEC ID n° 135 y SEC ID n° 137;
- 40 - una composición diagnóstica, profiláctica, o terapéutica, en particular para inhibir la expresión de por lo menos un retrovirus asociado a la esclerosis en placas y/o a la poliartritis reumatoide, que comprende un fragmento nucleotídico definido anteriormente;
- 45 - un procedimiento para detectar un retrovirus asociado a la esclerosis en placas y/o a la poliartritis reumatoide, en una muestra biológica, que comprende las etapas que consisten en poner en contacto un ARN y/o un ADN que pertenece presuntamente o que procede de dicho retrovirus, o sus ARN y/o ADN complementarios, con una composición que comprende un fragmento nucleotídico definido anteriormente.

50 Antes de detallar la invención, se definirán ahora diferentes términos usados en la descripción y en las reivindicaciones:

- 55 - por cepa o aislado, se entiende cualquier fracción biológica infectante y/o patógena, que contiene, por ejemplo, unos virus y/o unas bacterias y/o unos parásitos, que generan un poder patógeno y/o antigénico, hospedada por un cultivo o un hospedante vivo; a título de ejemplo, una cepa vírica según la definición anterior puede contener un agente co-infectante, por ejemplo una protista patógena,
- 60 - el término "MSRV" usado en la presente descripción designa cualquier agente patógeno y/o infectante, asociado a la SEP, en particular una especie vírica, las cepas atenuadas de dicha especie vírica, o las partículas defectivas interferentes o que contienen unos genomas co-encapsidados o también unos genomas recombinados con una parte del genoma MSRV-1, derivados de esta especie. Se sabe que los virus y particularmente los virus que contienen ARN, tienen una variabilidad consecutiva en particular de unos índices relativamente elevados de mutación espontánea, la cual se tendrá en cuenta a continuación para definir la noción de equivalencia,
- 65 - por virus humano, se entiende un virus susceptible de infectar o de ser hospedado por el ser humano,

ES 2 331 251 T3

- 5 - teniendo en cuenta todas las variaciones y/o recombinaciones naturales o inducidas que se pueden encontrar en la práctica de la presente invención, los objetos de esta última, definidos anteriormente y en las reivindicaciones, se han expresado comprendiendo los equivalentes o derivados de los diferentes materiales biológicos definidos a continuación, en particular de las secuencias homólogas nucleotídicas o peptídicas,
- 10 - la variante de un virus o de un agente patógeno y/o infectante según la invención, comprende por lo menos un antígeno reconocido por lo menos por un anticuerpo dirigido contra por lo menos un antígeno correspondiente de dicho virus y/o de dicho agente patógeno y/o infectante, y/o un genoma del que cualquier parte se detecta por lo menos por una sonda de hibridación, y/o por lo menos un cebador de amplificación nucleotídica específica de dicho virus y/o agente patógeno y/o infectante, en unas condiciones de hibridación determinadas bien conocidas por el experto en la materia,
- 15 - según la invención, un fragmento nucleotídico o un oligonucleótido o un polinucleótido es una cadena de monómeros, o un biopolímero, caracterizado por la secuencia informacional de los ácidos nucleicos naturales, susceptible de hibridarse a cualquier otro fragmento nucleotídico en unas condiciones predeterminadas, pudiendo contener la cadena unos monómeros de estructuras químicas diferentes y ser obtenida a partir de una molécula de ácido nucleico natural y/o por recombinación genética y/o por síntesis química; un fragmento nucleotídico puede ser idéntico a un fragmento genómico del virus MSRV-1 considerado por la presente invención, en particular un gen de este último, por ejemplo pol o env en el caso de dicho virus;
- 20 - así, un monómero puede ser un nucleótido natural de ácido nucleico, cuyos elementos constitutivos son un azúcar, un grupo fosfato y una base nitrogenada; en el ARN el azúcar es la ribosa, en el ADN el azúcar es la desoxi-2-ribosa; según que se trate del ADN o del ARN, la base nitrogenada se selecciona de entre la adenina, la guanina, el uracilo, la citosina, y la timina; o el nucleótido puede ser modificado en uno por lo menos de los tres elementos constitutivos; a título de ejemplo, la modificación puede intervenir a nivel de las bases, generando unas bases modificadas tales como la inosina, la metil-5-desoxicitidina, la desoxiuridina, la dimetilamino-5-desoxiuridina, la diamino-2,6-purina, la bromo-5-desoxiuridina y cualquier otra base modificada que favorece la hibridación; a nivel del azúcar, la modificación puede consistir en la sustitución de por lo menos una desoxiribosa por una poliamida, y a nivel del grupo fosfato, la modificación puede consistir en su sustitución por unos ésteres, en particular seleccionados de entre los ésteres de difosfato, de alquilo y arilfosfonato y de fosforotioato,
- 25 - mediante la expresión “secuencia informacional” se entiende cualquier sucesión ordenada de monómeros, cuya naturaleza química y el orden en un sentido de referencia, constituyen o no una información funcional de la misma calidad que la de los ácidos nucleicos naturales,
- 30 - por hibridación se entiende el proceso durante el cual, en unas condiciones de funcionamiento apropiadas, dos fragmentos nucleotídicos, que tienen unas secuencias suficientemente complementarias, se aparean para formar una estructura compleja, en particular doble o triple, preferentemente en forma de hélice,
- 35 - una sonda comprende un fragmento nucleotídico sintetizado por vía química u obtenido por digestión o corte enzimático de un fragmento nucleotídico más largo, que comprende por lo menos seis monómeros, ventajosamente de 10 a 100 monómeros, preferentemente de 10 a 30 monómeros, y que posee una especificidad de hibridación en unas condiciones determinadas; preferentemente, una sonda que comprende menos de 10 monómeros no se usa sola, sino en presencia de otras sondas de tamaño tan corto o no; en ciertas condiciones particulares, puede resultar útil usar unas sondas de tamaño superior a 100 monómeros; una sonda puede ser usada en particular con fines de diagnóstico y se tratará, por ejemplo, de sondas de captura y/o de detección,
- 40 - la sonda de captura se puede inmovilizar sobre un soporte sólido mediante cualquier medio apropiado, es decir, directa o indirectamente, por ejemplo por covalencia o adsorción pasiva,
- 45 - la sonda de detección se puede marcar mediante un marcador seleccionado en particular de entre los isótopos radioactivos, unas enzimas seleccionadas en particular de entre la peroxidasa y la fosfatasa alcalina y aquéllos susceptibles de hidrolizar un sustrato cromógeno, fluorígeno o luminiscente, unos compuestos químicos cromóforos, unos compuestos cromógenos, fluorígenos o luminiscentes, unos análogos de bases nucleotídicas, y la biotina,
- 50 - las sondas usadas con fines de diagnóstico de la invención se pueden usar en cualquier técnica de hibridación conocida, y en particular las técnicas denominadas “DOT-BLOT”, “SOUTHERN BLOT”, “NORTHERN BLOT” que es una técnica idéntica a la técnica “SOUTHERN BLOT” pero que usa ARN como diana, la técnica SANDWICH; ventajosamente, se usa la técnica SANDWICH en la presente invención, que comprende una sonda de captura específica y/o una sonda de detección específica, entendiéndose que la sonda de captura y la sonda de detección deben presentar una secuencia nucleotídica por lo menos parcialmente diferente,
- 55
- 60
- 65

ES 2 331 251 T3

- cualquier sonda según la presente invención puede hibridarse *in vivo* o *in vitro* sobre el ARN y/o sobre el ADN, para bloquear los fenómenos de replicación, en particular traducción y/o transcripción, y/o para degradar dicho ADN y/o ARN,
- 5 - un cebador es una sonda que comprende por lo menos seis monómeros, y ventajosamente de 10 a 30 monómeros, que poseen una especificidad de hibridación en unas condiciones determinadas, para la iniciación de una polimerización enzimática, por ejemplo en una técnica de amplificación tal como la PCR (Polymerase Chain Reaction), en un procedimiento de elongación, tal como la secuenciación, en un método de transcripción inversa o análogo,
- 10 - dos secuencias nucleotídicas o peptídicas son denominadas equivalentes o derivadas una con relación a la otra, o con relación a una secuencia de referencia, si funcionalmente los biopolímeros correspondientes pueden desempeñar sustancialmente la misma función, sin ser idénticas, frente a la aplicación o al uso considerado, o en la técnica en la que intervienen; en particular, son equivalentes dos secuencias obtenidas del hecho de la variabilidad natural, en particular mutación espontánea de la especie a partir de la cual se han identificado, o se han inducido, así como dos secuencias homólogas, siendo la homología definida a continuación,
- 15 - por “variabilidad” se entiende cualquier modificación, espontánea o inducida de una secuencia, en particular por sustitución, y/o inserción, y/o delección de nucleótidos y/o de fragmentos nucleotídicos, y/o extensión y/o acortamiento de la secuencia en uno por lo menos de los extremos; una variabilidad no natural puede resultar de las técnicas de ingeniería genética usadas, por ejemplo de la elección de los cebadores de síntesis degenerados o no, seleccionados para amplificar un ácido nucleico; esta variabilidad puede traducirse por unas modificaciones de cualquier secuencia de partida, considerada como referencia, y que pueden ser expresadas por un grado de homología con relación a dicha secuencia de referencia,
- 20 - la homología caracteriza el grado de identidad de dos fragmentos nucleotídicos o peptídicos comparados; se mide por el porcentaje de identidad que está determinado en particular por la comparación directa de secuencias nucleotídicas o peptídicas, con relación a unas secuencias nucleotídicas o peptídicas de referencia,
- 25 - cualquier fragmento nucleotídico se denomina equivalente o derivado de un fragmento de referencia si presenta una secuencia nucleotídica equivalente a la secuencia del fragmento de referencia; según la definición anterior, en particular son equivalentes a un fragmento nucleotídico de referencia:
 - 30 (a) cualquier fragmento susceptible de hibridarse por lo menos parcialmente con el complemento del fragmento de referencia,
 - (b) cualquier fragmento cuya alineación con el fragmento de referencia conduce a demostrar unas bases contiguas idénticas, en número más importante que con cualquier otro fragmento que procede de otro grupo taxonómico,
 - (c) cualquier fragmento que resulta o que puede resultar de la variabilidad natural de la especie, a partir de la cual se obtiene,
 - 40 (d) cualquier fragmento que puede resultar de las técnicas de ingeniería genética aplicadas al fragmento de referencia,
 - (e) cualquier fragmento, que comprende por lo menos ocho nucleótidos contiguos, que codifica para un péptido homólogo o idéntico al péptido codificado por el fragmento de referencia,
 - 45 (f) cualquier fragmento diferente del fragmento de referencia, mediante inserción, delección, sustitución de por lo menos un monómero, extensión, o acortamiento en uno por lo menos de sus extremos; por ejemplo, cualquier fragmento que corresponde al fragmento de referencia, flanqueado en uno por lo menos de sus extremos por una secuencia nucleotídica que no codifica para un polipéptido,
 - 50 - por polipéptido se entiende en particular cualquier péptido de por lo menos dos aminoácidos, en particular oligopéptido, proteína, extracto, separado, o sustancialmente aislado o sintetizado, por la intervención del hombre, específicamente los que se obtienen por síntesis química, o por expresión en un organismo recombinante,
 - 55 - por polipéptido codificado de manera parcial por un fragmento nucleotídico, se entiende un polipéptido que presenta por lo menos tres aminoácidos codificados por lo menos por nueve monómeros contiguos incluidos en dicho fragmento nucleotídico,
 - 60 - un aminoácido se denomina análogo a otro aminoácido, cuando sus características fisicoquímicas respectivas, tales como la polaridad, la hidrofobicidad, y/o la alcalinidad, y/o acidez, y/o neutralidad, son sustancialmente las mismas; de esta manera, una leucina es análoga a una isoleucina,
 - 65

ES 2 331 251 T3

- cualquier polipéptido se denomina equivalente o derivado de un polipéptido de referencia, si los polipéptidos comparados tienen sustancialmente las mismas propiedades, y en particular las mismas propiedades antigénicas, inmunológicas, enzimológicas y/o de reconocimiento molecular; en particular es equivalente a un polipéptido de referencia:

5

(a) cualquier polipéptido que comprende una secuencia de la que por lo menos un aminoácido ha sido sustituido por un aminoácido análogo,

10

(b) cualquier polipéptido que tiene una secuencia peptídica equivalente, obtenida por variación natural o inducida de dicho polipéptido de referencia, y/o del fragmento nucleotídico que codifica para dicho polipéptido,

(c) un mimotopo de dicho polipéptido de referencia,

15

(d) cualquier polipéptido en cuya secuencia se sustituyen uno o varios aminoácidos de la serie L por un aminoácido de la serie D, y viceversa,

20

(e) cualquier polipéptido en cuya secuencia se ha introducido una modificación de las cadenas laterales de los aminoácidos, tal como por ejemplo una acetilación de las funciones aminadas, una carboxilación de las funciones tiol, una esterificación de las funciones carboxílicas,

(f) cualquier polipéptido en cuya secuencia han sido modificados uno o varios enlaces peptídicos, como por ejemplo los enlaces carba, retro, inverso, retro-inverso, reducidos, y metilenoxi,

25

(g) cualquier polipéptido del que por lo menos un antígeno es reconocido por un anticuerpo dirigido contra un polipéptido de referencia,

- el porcentaje de identidad que caracteriza la homología de dos fragmentos peptídicos comparados es según la presente invención de por lo menos 50% y preferentemente, de por lo menos 70%.

30

Puesto que un virus que posee una actividad enzimática transcriptasa inversa puede ser genéticamente caracterizado tanto en forma de ARN como de ADN, se mencionará tanto el ADN como el ARN viral para caracterizar las secuencias relativas a un virus que posee dicha actividad transcriptasa inversa, denominado MSRV-1 según la presente descripción.

35

Las expresiones de orden usadas en la presente descripción y en las reivindicaciones, tales como “primera secuencia nucleotídica” no se consideran para expresar un orden particular, sino para definir más claramente la invención.

40

Mediante la expresión “detección de una sustancia o agente”, se entiende a continuación tanto una identificación como una cuantificación, o una separación o aislamiento de dicha sustancia o de dicho agente.

La invención se pondrá más claramente de manifiesto a partir de la lectura de la descripción detallada siguiente, realizada haciendo referencia a las figuras adjuntas, en las que:

45

La figura 1 representa la estructura general del ADN proviral y el ARN genómico de MSRV-1.

La figura 2 representa la secuencia nucleotídica del clon denominado CL6-5' (SEC ID n° 112) y tres marcos de lectura potenciales en aminoácidos que figuran bajo la secuencia nucleotídica.

50

La figura 3 representa la secuencia nucleotídica del clon denominado CL6-3' (SEC ID n° 114) y tres marcos de lectura potenciales en aminoácidos que figuran bajo la secuencia nucleotídica.

La figura 4 representa la secuencia nucleotídica del clon denominado C15 (SEC ID n° 117) y tres marcos de lectura potenciales en aminoácidos que figuran bajo la secuencia nucleotídica.

55

La figura 5 representa la secuencia nucleotídica del clon denominado 5M6 (SEC ID n° 120) y tres marcos de lectura potenciales en aminoácidos que figuran bajo la secuencia nucleotídica.

60

La figura 6 representa la secuencia nucleotídica del clon denominado CL2 (SEC ID n° 130) y tres marcos de lectura potenciales en aminoácidos que figuran bajo la secuencia nucleotídica.

La Figura 7 representa tres marcos de lectura potenciales en aminoácidos expresados por pET28C-clon 2 y que figuran bajo la secuencia nucleotídica.

65

La Figura 8 representa tres marcos de lectura potenciales en aminoácidos expresados por pET21C-clon 2 y que figuran bajo la secuencia nucleotídica.

ES 2 331 251 T3

La figura 9 representa la secuencia nucleotídica del clon denominado LB13 (SEC ID nº 141) y tres marcos de lectura potenciales en aminoácidos que figuran bajo la secuencia nucleotídica.

La figura 10 representa la secuencia nucleotídica del clon denominado LA15 (SEC ID nº 142) y tres marcos de lectura potenciales en aminoácidos que figuran bajo la secuencia nucleotídica.

La figura 11 representa la secuencia nucleotídica del clon denominado LB16 (SEC ID nº 124) y tres marcos de lectura potenciales en aminoácidos que figuran bajo la secuencia nucleotídica.

La figura 12 representa la actividad promotora expresada en cpm/4 min de las secuencias U3R subclonadas a partir de LTR de orígenes diferentes en el plásmido PCAT3. PCAT3 significa plásmido solo, PCAT-PH74 significa plásmido más clon U3R endógeno expresado en la placenta, PCAT-c16 significa plásmido más clon U3R amplificado en el ARN de un plasma SEP, PCAT-5M6 significa plásmido más región U3R amplificada en el ADN celular, “sin plásmido” significa la ausencia de plásmido en el ensayo.

La Figura 13 representa las secuencias MSRV1 env y LTR3'. Las flechas horizontales indican el inicio de las regiones env, U3 y R. En la región env: el péptido señal y la región inmunosupresiva potencial están subrayados, los sitios de glicosilación potenciales están enmarcados y los sitios de escisión potenciales se indican mediante unas flechas verticales. En la región U3R: el elemento de regulación CAAT y la TATA Box están subrayados, el sitio “cap” y la señal de poliadenilación también están indicados.

La Figura 14 representa una región LTR5' (RU5) seguida de un sitio PBS (primer binding site) complementario del ARNt Trp y de un gen gag que codifica para una proteína de aproximadamente 487 aminoácidos. Los aminoácidos conservados en la nucleocápsida están subrayados dos veces. Los aminoácidos que definen la región de mayor homología en la cápsida aparecen en negrita y están subrayados una vez. Las / en la secuencia de aminoácidos indican unas variaciones observadas según los clones y en la secuencia nucleotídica indican unos saltos de marco en ciertos clones. Las regiones enmarcadas corresponden a unos epítomos identificados por análisis peptídico de la región C-terminal.

La Figura 15 representa la región integrasa de MSRV1, la secuencia nucleotídica y la secuencia en aminoácidos deducida de la región integrasa que corresponde al clon 87-23. En la Figura 15 // significa un salto de marco que ha sido suprimido para restituir el ORF potencial. Las letras en negrita subrayadas representan los aminoácidos conservados de las integrasas retrovíricas.

La Figura 16 describe las secuencias nucleotídica y peptídica del clon B13 (idéntico al clon FBd13 descrito en unas solicitudes anteriores) con indicación de los ORF y de codones stop representados por un punto. La región subrayada y en negrita representa el dominio inmunosupresor potencial. El dominio solo subrayado representa el inicio del LTR3'.

Ejemplo 1

Obtención de una región CL6-5 que codifica para el extremo n-terminal de la integrasa y de una región CL6-3' que contiene la secuencia 3' terminal del genoma MSRV-1

Una 3'RACE ha sido efectuada sobre el ARN total extraído de plasma de un paciente que padece SEP. Un plasma de control sano, tratado en las mismas condiciones, ha sido usado como control negativo. La síntesis de ADNc ha sido realizada con un cebador oligo dT identificado por la SEC ID nº 68 (5' GAC TCG CTG CAG ATC GAT TTT TTT TTT TTT TTT T 3') y la transcriptasa inversa “Expand™ RT” de Boehringer según las condiciones recomendadas por la compañía: Una PCR ha sido efectuada con la enzima Klentaq (Clontech) bajo las siguientes condiciones: 94°C durante 5 min., después 93°C durante 1 min., 58°C durante 1 min., 68°C durante 3 min. durante 40 ciclos y 68°C durante 8 min., con un volumen de reacción final de 50 µl.

Cebadores usados para la PCR:

- cebador 5', identificado por la SEC ID nº69
5' GCC ATC AAG CCA CCC AAG AAC TCT TAA CTT 3'
- cebador 3', identificado por la SEC ID nº68

Una segunda PCR denominada “semi anidada” ha sido realizada con un cebador 5' situado en el interior de la región ya amplificada. Esta segunda PCR ha sido efectuada bajo las mismas condiciones experimentales que las usadas durante la primera PCR, utilizando 10 µl del producto de amplificación resultante de la primera PCR.

ES 2 331 251 T3

Cebadores usados para la PCR semi-anidada:

- cebador 5', identificado por la SEC ID n° 70

5' CCA ATA GCC AGA CCA TTA TAT ACA CTA ATT 3';

- cebador 3', identificado por la SEC ID n° 68

10 Los cebadores SEC ID n° 69 y SEC ID n° 70 son específicos de la región pol de MSRV-1.

Un producto de amplificación de 1,9 Kb ha sido obtenido para el plasma de paciente SEP. El fragmento correspondiente no ha sido observado en el plasma de control sano. Este producto de amplificación ha sido clonado de la manera siguiente:

15 el ADN amplificado ha sido insertado en un plásmido con la ayuda del kit TA Cloning®: Los 2 µl de disolución de ADN han sido mezclados con 5 µl de agua destilada estéril, 1 µl de un tampón de ligación 10 veces concentrado "10X LIGATION BUFFER", 2 µl de "pCR™ VECTOR" (25 ng/ml) y 1 µl de "T4 DNA LIGASA". Esta mezcla se incubó una noche a 12°C. Las etapas siguientes han sido realizadas de acuerdo con las instrucciones del kit TA Cloning® (Invitrogen). Al final del proceso, las colonias blancas de bacterias recombinantes (white) han sido inoculadas nuevamente para ser cultivadas y permitir la extracción de los plásmidos incorporados según el proceso denominado "miniprep". La preparación de plásmido de cada colonia recombinante ha sido cortada por una enzima de restricción apropiada y analizada sobre gel del agarosa. Los plásmidos que poseen un inserto detectado bajo luz UV después del marcado del gel con bromuro de etidio, han sido seleccionados para la secuenciación del inserto, después de la hibridación con un cebador complementario del promotor Sp6 presente sobre el plásmido del TA Cloning kit®. La reacción previa a la secuenciación ha sido efectuada a continuación según el método recomendado para el uso del kit de secuenciación "PRISM™ Ready Reaction AmpliTaq® FS, DyeDeoxy™ Terminator" (Applied Biosystems, ref. 402119) y la secuenciación automática ha sido realizada en los aparatos 373 A y 377 Applied Biosystems, según las instrucciones del fabricante.

El clon obtenido contiene una región CL6-5' que codifica para el extremo N terminal de la integrasa y una región CL6-3', que corresponde a la región 3' terminal de MSRV-1 y que permite definir el final de la envoltura (234 pb) y las regiones U3, R (401 pb) del retrovirus MSRV1.

La región que corresponde al extremo N terminal de la integrasa está representada por su secuencia nucleotídica (SEC ID n° 112) en la figura 27. Los tres marcos de lectura potenciales están presentados por su secuencia aminoácida bajo la secuencia nucleotídica, y la secuencia de aminoácidos del extremo N-terminal de la integrasa se identifica por la SEC ID n° 113.

La región C16-3' está representada por su secuencia nucleotídica (SEC ID n° 114) en la figura 3. Los tres marcos de lectura potenciales están presentados por su secuencia aminoácida bajo la secuencia nucleotídica. Una secuencia aminoácida que corresponde al extremo C-terminal de la proteína env de MSRV-1 se identifica por la SEC ID n° 115.

Con el fin de evaluar la actividad promotora del LTR obtenido a partir del clon 6 (c16), se ha realizado un ensayo de actividad promotora que usa la enzima CAT (cloramfenicol acetil transferasa) con la región U3R correspondiente. Paralelamente, se ha ensayado un clon que contiene la misma región U3R de ARN retroviral endógeno expresado en la placenta normal (PH74) y un clon (5M6) que procede de ADN. El resultado presentado en la figura 12 muestra una acentuada actividad promotora del LTR que procede de plasma SEP (c16) y una actividad significativamente más débil con las secuencias de origen endógeno no SEP.

Ejemplo 2

55 *Obtención del clon C15 que contiene la región que codifica para una parte de la envoltura del retrovirus MSRV-1*

Una RT-PCR ha sido efectuada sobre el ARN total extraído de viriones concentrados por ultracentrifugación a partir de un sobrenadante de cultivo de sinoviocitos que proceden de un paciente PR. La síntesis de ADNc ha sido realizada con un cebador oligo dT y la transcriptasa inversa "Expand™ RT" de Boehringer según las condiciones recomendadas por la compañía. Una PCR ha sido efectuada con el Expand™ Long Template PCR System (Boehringer) bajo las condiciones siguientes: 94°C durante 5 min. y después 93°C durante 1 min., 60°C durante 1 min., 68°C durante 3 min. durante 40 ciclos y 68°C durante 8 min. y con un volumen de reacción final de 50 µl.

Cebadores usados para la PCR:

- cebador 5', identificado por la SEC ID n°: 69

5' GCC ATC AAG CCA CCC AAG AAC TCT TAA CTT 3';

ES 2 331 251 T3

- cebador 3', identificado por la SEC ID nº 116

5' TGG GGT TCC ATT TGT AAG ACC ATC TGT AGC TT 3'

5 Una segunda PCR denominada "semi-anidada" ha sido realizada con un cebador 5' situado en el interior de la región ya amplificada. Esta segunda PCR ha sido efectuada bajo las mismas condiciones experimentales que las usadas durante la primera PCR (salvo que se realizaron 30 ciclos en vez de 40), usando 10 µl del producto de amplificación resultante de la primera PCR.

10 Cebadores usados para la PCR semi-anidada:

- cebador 5', identificado por la SEC ID nº70

5' CCA ATA GCC AGA CCA TTA TAT ACA CTA ATT 3'

15

- cebador 3', identificado por la SEC ID nº 116

20 Los cebadores SEC ID nº 69 y SEC ID nº 70 son específicos de la región pol de MSRV-1. El cebador SEC ID nº 116 es específico de la secuencia FBd13 (también denominada B13) y está localizado en la región env conservada entre los oncoretrovirus.

Un producto de amplificación de 1932 pb ha sido obtenido y clonado de la manera siguiente:

25 el ADN amplificado ha sido insertado en un plásmido con la ayuda del kit TA Cloning®. Las diferentes etapas han sido realizadas de acuerdo con las instrucciones del kit TA Cloning® (Invitrogen). Al fin del proceso, las colonias blancas de bacterias recombinantes (white) han sido inoculadas nuevamente para ser cultivadas y permitir la extracción de los plásmidos incorporados según el procedimiento denominado "miniprep". La preparación de plásmido de cada colonia recombinante ha sido cortada por una enzima de restricción apropiada y analizada sobre gel del agarosa. Los plásmidos que poseen un inserto detectado bajo luz UV después del marcado del gel con bromuro de etidio, han sido seleccionados para la secuenciación del inserto, después de la hibridación con un cebador complementario del promotor SP6 presente sobre el plásmido de clonación del TA cloning kit®. La reacción previa a la secuenciación ha sido efectuada a continuación según el método recomendado para el uso del kit de secuenciación "PRISM™ Ready Reaction AmpliTaq® FS, DyeDeoxy™ Terminator" (Applied Biosystems, ref. 402119) y la secuenciación automática ha sido realizada en los aparatos 373 A y 377 de Applied Biosystems, según las instrucciones del fabricante.

40 El clon C15 obtenido contiene una región que corresponde a la región de envoltente de MSRV-1, de 1.481 pb.

45 La región env del clon C15 está representada por su secuencia nucleotídica (SEC ID nº 117) en la figura 5. Los tres marcos de lectura potenciales de este clon están presentados por su secuencia aminoácida bajo la secuencia nucleotídica. El marco de lectura que corresponde a una proteína env estructural MSRV-1 se identifica por la SEC ID nº 118.

A partir de las secuencias definidas que proceden de los clones c16 y C15, se ha podido efectuar una construcción plasmídica que codifica para una envoltente completa seguida del LTR3', tal como se presenta en la figura 13 con el marco de lectura correspondiente.

50 Ejemplo 3

Obtención de un clon 5M6 que contiene las secuencias de la región 3' terminal de la envoltente, seguidas por las secuencias U3, R, U5 de tipo proviral MSRV-1

55 Una PCR monodireccional ha sido efectuada sobre el ADN extraído de linfocitos B inmortalizados en cultivo de un paciente PR. La PCR ha sido efectuada con Expand™ Long Template PCR System (Boehringer) bajo las condiciones siguientes: 94°C durante 3 min. y después 93°C durante 1 min., 60°C durante 1 min., 68°C durante 3 min. durante 10 ciclos, después 93°C durante 1 min., 60°C durante 1 min. con 15 s de extensión por cada ciclo, 68°C durante 3 min. durante 35 ciclos y 68°C durante 7 min. y con un volumen de reacción final de 50 µl.

60 El cebador usado para la PCR identificado por la SEC ID nº 119 es 5' TCA AAA TCG AAG AGC TTT AGA CTT GCT AAC CG 3';

65 El cebador SEC ID nº 119 es específico de la región env del Clon C15.

ES 2 331 251 T3

Un producto de amplificación de 1.673 pb ha sido obtenido y clonado de la manera siguiente:

el ADN amplificado ha sido insertado en un plásmido con la ayuda del kit TA Cloning®. Las diferentes etapas han sido realizadas de acuerdo con las instrucciones del kit TA Cloning® (Invitrogen). Al final del proceso, las colonias blancas de bacterias recombinantes (white) han sido inoculadas nuevamente para ser cultivadas y permitir la extracción de los plásmidos incorporados según el procedimiento denominado “miniprep”. La preparación de plásmido de cada colonia recombinante ha sido cortada por una enzima de restricción apropiada y analizada sobre gel de agarosa. Los plásmidos que poseen un inserto detectado bajo luz UV después del marcado del gel con bromuro de etidio, han sido seleccionados para la secuenciación del inserto, después de la hibridación con un cebador complementario del promotor T7 presente sobre el plásmido de clonación del TA cloning kit®. La reacción previa a la secuenciación ha sido efectuada según el método recomendado para el uso del kit de secuenciación “PRISM™ Ready Reaction AmpliTaq® FS, DyeDeoxy™ Terminator” (Applied Biosystems, ref. 402119) y la secuenciación automática ha sido realizada en los aparatos 373 A y 377 Applied Biosystems, según las instrucciones del fabricante.

El clon 5M6 obtenido contiene una región que corresponde a la región 3' de la envoltura de MSRV-1, de 492 pb seguida de las regiones U3, R y U5 (837 pb) de MSRV-1.

El clon 5M6 está representado por su secuencia nucleotídica (SEC ID n° 120) en la figura 5. Los tres marcos de lectura potenciales de este clon se representan por su secuencia aminoácida bajo la secuencia nucleotídica. El marco de lectura que corresponde al extremo C-terminal de la proteína env MSRV-1 se identifica por la SEC ID n° 121.

Ejemplo 4

Obtención del clon LB16 que contiene la región que codifica la integrasa del retrovirus MSRV-1

Una RT-PCR ha sido efectuada sobre el ADN total tratado con la DNAsaI y extraído a partir de un plexo coroideo que procede de un paciente que padece SEP. La síntesis de ADNc ha sido realizada con un cebador oligo dT y la transcriptasa inversa “Expand™ RT” de Boehringer según las condiciones recomendadas por la compañía. Un control “no RT” se ha efectuado paralelamente sobre el mismo material. Una PCR ha sido efectuada con la Taq polimerasa (Perkin Elmer) bajo las condiciones siguientes: 95°C durante 5 min., y 95°C durante 1 min., 55°C durante 1 min., 72°C durante 2 min. durante 35 ciclos y 72°C durante 8 min. y con un volumen de reacción final de 50 µl.

Cebadores usados para la PCR:

- cebador 5', identificado por la SEC ID n° 122

5' GGC ATT GAT AGC ACC CAT CAG 3';

- cebador 3', identificado por la SEC ID n° 123

5' CAT GTC ACC AGG GTG GAA TAG 3'

El cebador SEC ID n° 122 es específico de la región pol de MSRV-1 y más precisamente similar a la región integrasa descrita anteriormente. El cebador SEC ID n° 123 ha sido definido sobre unas secuencias de los clones obtenidos durante ensayos anteriores.

Un producto de amplificación de aproximadamente 760 pb ha sido obtenido únicamente en el ensayo con RT y ha sido clonado de la manera siguiente:

el ADN amplificado ha sido insertado en un plásmido con la ayuda del kit TA Cloning®. Las diferentes etapas han sido realizadas según las instrucciones del kit TA Cloning® (Invitrogen). Al final del proceso, las colonias blancas de bacterias recombinantes (white) han sido inoculadas nuevamente para ser cultivadas y permitir la extracción de los plásmidos incorporados según el procedimiento denominado “miniprep”. La preparación de plásmido de cada colonia recombinante ha sido cortada por una enzima de restricción apropiada y analizada sobre gel de agarosa. Los plásmidos que poseen un inserto detectado bajo luz UV después del marcado del gel con bromuro de etidio, han sido seleccionados para la secuenciación del inserto, después de la hibridación con un cebador complementario del promotor T7 presente sobre el plásmido de clonación del TA cloning kit®. La reacción previa a la secuenciación ha sido efectuada a continuación según el método recomendado para el uso del kit de secuenciación “PRISM™ Ready Reaction AmpliTaq® FS, DyeDeoxy™ Terminator” (Applied Biosystems, ref. 402119) y la secuenciación automática ha sido realizada en los aparatos 373 A y 377 de Applied Biosystems, según las instrucciones del fabricante.

El clon LB16 obtenido contiene las secuencias que corresponden a la integrasa. La secuencia nucleotídica de este clon es identificada por la SEC ID n° 124 en la figura 11, estando determinados tres marcos de lectura.

ES 2 331 251 T3

Ejemplo 5

Obtención de un clon 2, CL2, que contiene en 3' una parte homóloga al gen *POL*, que corresponde al gen proteasa, y al gen *GAG* (*GM3*) que corresponde a la nucleocápsida, y una nueva región 5' codificante, que corresponde al gen *GAG* más específicamente la matriz y la cápsida de *MSRV-1*

Una amplificación por PCR ha sido efectuada sobre el ADN total extraído a partir de 100 μ l de plasma de un paciente que padece SEP. Un control con agua, tratado bajo las mismas condiciones, ha sido utilizado como control negativo. La síntesis de ADNc ha sido realizada con 300 pmoles de un cebador aleatorio (GIBCO-BRL, Francia) y la transcriptasa inversa "Expand RT" (BOEHRINGER MANNHEIM, Francia) según las condiciones recomendadas por la compañía. Una amplificación por PCR ("polymerase chain reaction") ha sido efectuada con la enzima Taq polimerasa (Perkin Elmer, Francia) usando 10 μ l de ADNc bajo las condiciones siguientes: 94°C durante 2 min., 55°C durante 1 min., 72°C durante 2 min. y 94°C durante 1 min., 55°C durante 1 min., 72°C durante 2 min. durante 30 ciclos y 72°C durante 7 min. y con un volumen de reacción final de 50 μ l.

Cebadores usados para la amplificación por PCR:

- cebador 5', identificado por la SEC ID N° 126

5' CGG ACA TCC AAA GTG ATG GGA AAC G 3';

- cebador 3', identificado por la SEC ID n° 127

5' GGA CAG GAA AGT AAG ACT GAG AAG GC 3'

Una segunda amplificación por PCR denominada "semi-anidada" ha sido realizada con un cebador 5' situado en el interior de la región ya amplificada. Esta segunda PCR ha sido efectuada bajo las mismas condiciones experimentales que las usadas durante la primera PCR, usando 10 μ l del producto de amplificación resultante de la primera PCR.

Cebadores usados para la amplificación por PCR semi-anidada:

- cebador 5', identificado por la SEC ID n° 128

5' CCT AGA ACG TAT TCT GUA GAA TTG GG 3';

- cebador 3', identificado por la SEC ID n° 129

5' TGG CTC TCA ATG GTC AAA CAT ACC CG 3'

Los cebadores SEC ID n° 128 y SEC ID n° 129 son específicos de la región pol, clon G+E+A, más específicamente la región E: posición de nucleótidos n° 423 a n° 448. Los cebadores usados en la región 5' han sido definidos sobre unas secuencias de clones obtenidos durante ensayos previos.

Un producto de amplificación de 1.511 pb ha sido obtenido a partir del ARN extraído del plasma de paciente SEP. El fragmento correspondiente no se observó para el control con agua. Este producto de amplificación ha sido clonado de la manera siguiente.

El ADN amplificado ha sido insertado en un plásmido con la ayuda del kit TA Cloning™. Los 2 μ l de disolución de ADN han sido mezclados con 5 μ l de agua destilada estéril, 1 μ l de un tampón de ligación 10 veces concentrado "10X LIGATION BUFFER", 2 μ l de "pCR™ VECTOR" (25 ng/ml) y 1 μ l de "T4 DNA LIGASA". Esta mezcla se incubó durante una noche a 14°C. Las etapas siguientes han sido realizadas según las instrucciones del kit TA Cloning® (Invitrogen). La mezcla se extendió después de la transformación de la ligación en unas bacterias *E. coli* INV α F'. Al final del proceso, las colonias blancas de bacterias recombinantes han sido inoculadas nuevamente para ser cultivadas y permitir la extracción de los plásmidos incorporados según el procedimiento denominado "minipreparación de ADN" (17). La preparación de plásmido de cada colonia recombinante ha sido cortada por la enzima de restricción *Eco* RI y analizada sobre gel de agarosa. Los plásmidos que poseen un inserto detectado bajo luz UV después del marcado del gel con bromuro de etidio, han sido seleccionados para la secuenciación del inserto, después de la hibridación con un cebador complementario del promotor T7 presente sobre el plásmido de clonación del TA cloning kit®. La reacción previa a la secuenciación ha sido efectuada a continuación según el método recomendado para el uso del kit de secuenciación "PRISM™ Ready Reaction AmpliTaq® FS, DyeDeoxy™ Terminator" (Applied Biosystems, ref. 402119) y la secuenciación automática ha sido realizada en los aparatos 373 A y 377 de Applied Biosystems, según las instrucciones del fabricante.

El clon obtenido, denominado CL2, contiene una región C-terminal similar a la región 5' terminal de los clones G+E+A de *MSRV-1*, que permite definir la región C-terminal del gen *gag* y una nueva región que corresponde a la región N-terminal del gen *gag* *MSRV-1*.

ES 2 331 251 T3

CL2 permite definir una región de 1.511 pb que presenta una fase abierta de lectura dentro de la región N-terminal de 1.077 pb que codifica para 359 aminoácidos y una fase no abierta de lectura, de 454 pb, que corresponde a la región C-terminal del gen gag MSRV-1.

5 La secuencia nucleotídica de CL2 se identifica por la SEC ID n° 130. Está representada en la figura 6, con los marcos de lectura potenciales en aminoácido.

10 El fragmento de 1.077 pb de CL2 que codifica para 359 aminoácidos ha sido amplificado por PCR con la enzima *Pwo* (5 U/ μ l) (Boehringer Mannheim, Francia) usando 1 μ l de la minipreparación del ADN del clon 2 bajo las condiciones siguientes: 95°C durante 1 min., 60°C durante 1 min., 72°C durante 2 min. durante 25 ciclos y con un volumen de reacción final de 50 μ l con la ayuda de los cebadores:

- cebador 5' (*Bam* HI), identificado por la SEC ID n° 132

15 5' TGC TGG AAT TCG GGA TCC TAG AAC GTA TTC 3' (30 mer), y

- cebador 3' (*Hind* III), identificado por la SEC ID n° 133

20 5' AGT TCT GCT CCG AAG CTT AGG CAG ACT TTT 3' (30 mer)

que corresponden, respectivamente, a la secuencia nucleotídica del clon 2 en posición -9 a 21 y 1066 a 1095.

25 El fragmento obtenido después de la PCR ha sido linearizado por *Bam* HI y *Hind* III y subclonado en los vectores de expresión pET28C y pET21C (NOVAGEN) linearizado por *Bam* HI y *Hind* III. La secuenciación del ADN del fragmento de 1.077 pb del clon 2 en los dos vectores de expresión ha sido realizado según el método recomendado para el uso del kit de secuenciación "PRISM™ Ready Reaction Amplitaq® FS, DyeDeoxy™ Terminator" (Applied Biosystems, ref. 402119) y la secuenciación automática ha sido realizada en los aparatos 373 A y 377 de Applied Biosystems, según las instrucciones del fabricante.

30 La expresión de la secuencia nucleotídica del fragmento de 1.077 pb del clon 2 por los vectores de expresión pET28C y pET21C se identifican respectivamente por las SEC ID n° 135 y SEC ID n° 137.

35 Ejemplo 6

Expresión del clon 2 EN Escherichia coli

40 Las construcciones pET28C-clon 2 (1.077 pb) y pET21C-clon 2 (1.077 pb) sintetizan, en la cepa bacteriana BL21 (DE3), una proteína en fusión N- y C-terminal para el vector pET28C y C-terminal para el vector pET21C con 6 Histidinas, de masa molecular aparente de aproximadamente 45 kDa, demostrada por electroforesis sobre gel de poliacrilamida SDS-PAGE (SDS = Dodecil Sulfato de Sodio) (Laemmli, 1970 (1)). La reactividad de la proteína ha sido demostrada frente a un anticuerpo monoclonal anti-histidina (DIANOVA) mediante la técnica de transferencia Western (Towbin *et al.*, 1979 (2)).

45 Las proteínas recombinantes pET28C-clon 2 (1.077 pb) y pET21C-clon 2 (1.077 pb) han sido visualizadas en SDS-PAGE en la fracción insoluble después de la digestión enzimática de los extractos bacterianos con 50 μ l de lisozima (10 mg/ml) y lisis por ultrasonidos.

50 Las propiedades antigénicas de los antígenos recombinantes pET28C-clon 2 (1.077 pb) y pET21C-clon 2 (1.077 pb) han sido ensayados mediante transferencia Western después de la solubilización del residuo bacteriano con SDS al 2% y β -mercaptoetanol 50 mM. Después de la incubación con los sueros de pacientes que padecen esclerosis en placas, los sueros de los controles neurológicos y los sueros de los controles de centro de transfusión sanguínea (CTS), los inmunocomplejos han sido detectados con la ayuda de un suero de cabra anti-IgG y anti-IgM humanos, acoplado a la fosfatasa alcalina.

55 Los resultados se presentan en la tabla siguiente.

60

65

ES 2 331 251 T3

TABLA

Reactividad de sueros afectados por la esclerosis en placas y controles con la proteína recombinante MSR-1 gag clon 2 (1.077 pb) = pET21C-clon 2 (1.077 pb) y pET28c-clon 2 (1.077 pb)^a

5

10

15

20

25

30

35

40

45

50

55

60

65

ENFERMEDAD	NÚMERO DE INDIVIDUOS ENSAYADOS	NÚMERO DE INDIVIDUOS POSITIVOS
SEP	15	6 2 (+++), 2 (++), 2 (+)
CONTROLES NEUROLÓGICOS	2	1 (+++)
CONTROLES SANOS (CTS)	22	1 (+/-)

(a) Las tiras que contienen 1,5 µg de antígeno recombinante pET-gag clon 2 (1.077 pb) presentan una reactividad contra sueros diluidos al 1/100. La interpretación de Western Blot se basa en la presencia o en la ausencia de una banda pET-gag clon 2 (1.077 pb) específica sobre las tiras. Controles positivos y negativos están incluidos en cada experimento.

Estos resultados muestran que, en las condiciones técnicas usadas, aproximadamente 40% de los sueros humanos afectados de esclerosis en placas ensayados reaccionan con las proteínas recombinantes pET28C-clon 2 (1.077 pb) y pET21C-clon 2 (1.077 pb). Se observó una reactividad en un control neurológico y es interesante apuntar que los ARN extraídos a partir de este suero, después de la etapa de transcriptasa inversa, se amplifican asimismo por PCR en la región pol. Esto sugiere que unas personas que no han desarrollado una SEP pueden asimismo hospedar y expresar este virus. Por el contrario, un control (donante CTS) aparentemente sano, posee unos anticuerpos anti-gag (clon 2, 1.077 pb). Lo cual es compatible con una inmunidad adquirida contra MRSV-1 fuera de una enfermedad autoinmune asociada declarada.

Ejemplo 7

Obtención de un clon LB13 que contiene en 3' una parte homóloga al clon 2 que corresponde al gen GAG y en 5' una parte homóloga al clon 5M6 que corresponde a la región LTR U5

Una RT-PCR ("reverse-transcriptase-polymerase chain reaction") ha sido efectuada a partir del ARN total extraído de viriones, que proceden de sobrenadantes de células linfocitarias B de los pacientes que padecen esclerosis en placas, concentrados mediante ultracentrifugaciones. La síntesis de ADNc ha sido realizada con un cebador específico SEC n° XXX y la transcriptasa inversa "Expand™ RT" de BOEHRINGER MANNHEIM según las condiciones recomendadas por la compañía.

Cebador usado para la síntesis del ADNc, identificado por la SEC ID n° 138:

5' CTT GGA GGG TGC ATA ACC AGG GAA T 3'

Una amplificación por PCR ha sido realizada con la *Taq* polimerasa (perkin Elmer, Francia) en las condiciones siguientes: 94°C durante 1 min., 55°C durante 1 min., 72°C durante 2 min. durante 35 ciclos y 72°C durante 7 min. y con un volumen de reacción final de 100 µl.

Cebadores usados para la amplificación por PCR:

- cebador 5', identificado por la SEC ID n° 139

5' TGT CCG CTG TGC TCC TGA TC 3'

- cebador 3', identificado por la SEC ID n° 138

5' CTT GGA GGG TGC ATA ACC AGG GAA T 3'

Una segunda amplificación por PCR denominada "semi-anidada" ha sido realizada con un cebador 3' situado en el interior de la región ya amplificada. Esta segunda amplificación ha sido realizada bajo las mismas condiciones experimentales que las que se usaron durante la primera amplificación, usando 10 µl del producto de amplificación procedente de la primera PCR.

ES 2 331 251 T3

Cebadores usados para la amplificación por PCR “semi-anidada”:

- cebador 5', identificado por la SEC ID n° 139

5' TGT CCG CTG TGC TCC TGA TC 3'

- cebador 3', identificado por la SEC ID n° 140

5' CTA TGT CCT TTT GGA CTG TTT GGG T 3'

Los cebadores SEC ID n° 138 y SEC ID n° 140 son específicos de la región gag, clon 2 posición nucleotídica n° 373-397 y n° 433-456. Los cebadores usados en la región 5' han sido definidos sobre unas secuencias de los clones obtenidos durante unos ensayos previos.

Un producto de amplificación de 764 pb ha sido obtenido y clonado de la manera siguiente:

el ADN amplificado ha sido insertado en un plásmido con la ayuda del kit TA Cloning™. Los 2 µl de disolución de ADN han sido mezclados con 5 µl de agua destilada estéril, 1 µl de un tampón de ligación 10 veces concentrado “10X LIGATION BUFFER”, 2 µl de “pCR™ VECTOR” (25 ng/ml) y 1 µl de “T4 DNA LIGASE”. Esta mezcla se incubó durante una noche a 14°C. Las etapas siguientes han sido realizadas según las instrucciones del kit TA Cloning® (Invitrogen). Esta mezcla se extendió después de la transformación de la ligación en unas bacterias *E. coli* INVαF'. Al final del proceso, las colonias blancas de bacterias recombinantes han sido inoculadas nuevamente para ser cultivadas y permitir la extracción de los plásmidos incorporados según el procedimiento denominado “minipreparación de ADN” (17). La preparación de plásmido de cada colonia recombinante ha sido cortada por la enzima de restricción *Eco* RI y analizada sobre gel de agarosa. Los plásmidos que poseen un inserto detectado bajo luz UV después del marcado del gel con bromuro de etidio, han sido seleccionados para la secuenciación del inserto, después de la hibridación con un cebador complementario del promotor T7 presente sobre el plásmido de clonación del TA cloning kit®. La reacción previa a la secuenciación ha sido efectuada a continuación según el método recomendado para el uso del kit de secuenciación “PRISM™ Ready Reaction AmpliTaq® FS, DyeDeoxy™ Terminator” (Applied Biosystems, ref. 402119) y la secuenciación automática ha sido realizada en los aparatos 373 A y 377 de Applied Biosystems, según las instrucciones del fabricante.

El clon LB13 obtenido contiene una región N-terminal del gen gag MSRV-1 homóloga al clon 2 y un LTR que corresponde a una parte de la región U5. Entre la región U5 y gag se ha identificado un sitio de fijación para los ARN de transferencia, el PBS “primer binding site”.

La secuencia nucleotídica del fragmento de 764 pb del clon LB13 en el plásmido “pCR™ VECTOR” está representado en el identificador SEC ID n° 141.

El sitio de fijación para los ARN de transferencia, que presentan una secuencia de tipo PBS triptófano, ha sido identificado en posición nucleotídica n° 342-359 del clon LB13.

Puesto que este mismo PBS se ha encontrado nuevamente en las copias endógenas homólogas a MSRV-1, la familia endógena así definida se denominará de ahora en adelante HERV W, según la nomenclatura propuesta para las familias de retrovirus endógenos (W=Triptófano).

Un ORF corto de aproximadamente 65 aminoácidos se ha encontrado en la región U5 del LTR 5' del clon LB13.

Secuencia del ORF:

PMASNRAITLTAWSKIPFLGIRETKNPRSENTRLATMLEAAHHHEGSSPPLSWELWEQGPQVTIW.

La secuencia nucleotídica correspondiente que empieza con un codón ATG es susceptible de ser expresada en un ADN subgenómico a partir de un LTR proviral (U3RU5).

Otro clon, denominado LA15, ha sido obtenido sobre el ARN total extraído de viriones concentrados mediante ultracentrifugación a partir de un sobrenadante de cultivo de sinoviocitos que proceden de un paciente que padece poliartritis reumatoide. La estrategia de amplificación y clonación del clon LA15 es exactamente la misma que la que se ha usado para el clon LB13.

La secuencia nucleotídica del clon LA15 que está representada en el identificador SEC ID n° 142 es muy similar al clon LB13. Esto sugiere que el retrovirus MSRV-1 detectado en la esclerosis en placas presenta unas secuencias similares a las encontradas en la poliartritis reumatoide.

ES 2 331 251 T3

Ejemplo 8

Reconstrucción de una región RU5-GAG a partir de los clones LB15, LB13, CL2 y CL17

5 Los clones CL2 y LB13 ya se han descrito en los ejemplos anteriores. El clon LB15 se ha obtenido usando la secuencia R del LTR del clon c16 para definir un cebador en 5' y los cebadores anti-sentido usados son los mismos que para el clon LB13. El clon CL17 se ha obtenido por RT-PCR anidada usando los cebadores siguientes:

10 5'-TCATGCAACTGCACTCTTCTGGTCCG-3' (sentido)

5'-TCTTGCACCTAACCTCCACTGTCCGTTCC-3' (anti-sentido)

5'-ATCCCCCAGTAACAATTTGGTGACCACG-3' (sentido)

15 5'-TCGGGTCTAAGAGGGTACTTCCTTTGGTACC-3' (anti-sentido)

El clon LB15 se ha obtenido a partir de viriones obtenidos mediante cultivo de células de SEP. El clon LB17 se ha obtenido a partir del cultivo de plasma de paciente SEP.

20 Estos clones que se solapan han permitido reconstruir una secuencia RU5-gag con un ORF potencial dentro del gen GAG, tal como las representadas en la figura 14.

Ejemplo 9

Obtención de un clon 87-23

La región que corresponde a la integrasa se ha amplificado y clonado a partir de plasma de SEP usando una RT-PCR semi-anidada con los cebadores siguientes situados en las regiones pol y env de MSRV1.

30 En la región pol:

5'-TTACGCAGGTCTCAGGGATGAGCTT-3' (sentido-PCR primaria)

35 5'-CGGCAGTAGCAGTCTTAGTATCTGAAGCAGTTA-3' (sentido-PCR secundaria)

En la región env:

40 5'-GGTACGGAGGGTTTCATGTAGTTTTGAG-3' (PCR anti sentido primaria y secundaria)

El clon amplificado comprende 774 pb en la región pol/RT, toda la región integrasa (1.197 pb) y el inicio de la región env (480 pb). En la figura 15 se muestra la secuencia nucleotídica que corresponde a la región integrasa y la traducción en aminoácidos del ORF potencial.

Ejemplo 10

Confirmación de la presencia de ARN que contiene unas secuencias ENV parecidas a ERV9 en las partículas retrovirales asociadas al genoma MSRV1

55 Unas secuencias parecidas a ERV9 se han encontrado en proporción minoritaria en las preparaciones del virión procedente de SEP en relación con las secuencias MSRV1. Se ha descrito la existencia de fenómenos de co-encapsidación de secuencias endógenas filogenéticamente cercanas en las partículas retrovirales producidas por una cepa replicativa. De manera sorprendente, una región de ARN que comprende un inicio ORF en la parte 3' de env y que continúa potencialmente en el LTR3' ha sido encontrada de nuevo en diferentes muestras de SEP. Con el fin de precisar la existencia de un ORF se han realizado unos ensayos de transcripción-traducción y han permitido mostrar la realidad de un ORF env que contiene la totalidad de la parte transmembranaria (TM) y que se termina en el inicio del LTR putativo. Sin embargo, una trama adicional (ORFX) sigue y continúa dentro del LTR3'. Los dos productos de expresión se han visualizado y sus ORF respectivos han sido subclonados. La figura 16 representa las secuencias nucleotídica y peptídica del clon B13 descrito anteriormente precisando los ORF en la región env truncada y en el LTR putativo. La presencia de dichos ARN puede ser el origen de recombinaciones con la cepa replicativa, y por consiguiente, generar unas cepas de patogenicidad modificada.

65 Bibliografía

(1) Laemmli U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*. (1970). 227: 680-685.

ES 2 331 251 T3

(2) **Towbin H., Staehelin T. y Gordon J.** Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl. Acad. Sci. USA.* (1979). 76: 4350-4354.

5

10

15

20

25

30

35

40

45

50

55

60

65

REIVINDICACIONES

- 5 1. Material nucleico, en estado aislado o purificado, que comprende una secuencia nucleotídica seleccionada de entre el grupo constituido por (i) SEC ID n° 130; (ii) la secuencia complementaria de (i); y (iii) las secuencias que presentan para cualquier sucesión de 100 monómeros contiguos, por lo menos 50% de identidad con las secuencias (i) o (ii) respectivamente, siendo dicho material nucleico diferente de la secuencia HSAC64 del clon humano RG083M05 accesible en la base de datos EMBL, con el número AC000064.1.
- 10 2. Material nucleico según la reivindicación 1, **caracterizado** porque las secuencias (iii) son las secuencias que presentan para cualquier sucesión de 100 monómeros contiguos por lo menos 70% de identidad con las secuencias (i) o (ii) respectivamente.
- 15 3. Material nucleico, en estado aislado o purificado, que codifica para un polipéptido que presenta, para cualquier sucesión contigua de por lo menos 30 aminoácidos, por lo menos 50% de identidad con una secuencia peptídica seleccionada de entre el grupo constituido por la SEC ID n° 135 y SEC ID n° 137, siendo dicho material nucleico diferente de la secuencia HSAC64 del clon humano RG083M05 accesible en la base de datos EMBL, con el número AC000064.1.
- 20 4. Material nucleico según la reivindicación 3, **caracterizado** porque codifica para un polipéptido que presenta, para cualquier sucesión de por lo menos 30 aminoácidos, por lo menos 70% de identidad con SEC ID n° 135 o SEC ID n° 137.
- 25 5. Material nucleico retroviral, cuyo extremo 3' del gen gag termina en el nucleótido 1418 de SEC ID n° 130.
6. Material nucleico retroviral según cualquiera de las reivindicaciones anteriores, **caracterizado** porque está asociado a por lo menos una enfermedad autoinmune tal como la esclerosis en placas o la poliartritis reumatoide.
- 30 7. Fragmento nucleotídico que comprende una secuencia nucleotídica seleccionada de entre el grupo constituido por (i) SEC ID n° 130; (ii) la secuencia complementaria de (i); y (iii) las secuencias que presentan para cualquier sucesión de 100 monómeros contiguos, por lo menos 50%, y preferentemente por lo menos 70% de homología con las secuencias (i) o (ii) respectivamente, siendo dicho fragmento nucleotídico diferente de la secuencia HSAC64 del clon humano RG083M05 accesible en la base de datos EMBL, con el número AC000064.1.
- 35 8. Fragmento nucleotídico según la reivindicación 7, que consiste en una secuencia nucleotídica seleccionada de entre el grupo constituido por (i) SEC ID n° 130; (ii) la secuencia complementaria de (i); y (iii) las secuencias que presentan para cualquier sucesión de 100 monómeros contiguos, por lo menos 50% de identidad con las secuencias (i) o (ii) respectivamente.
- 40 9. Fragmento según la reivindicación 7 u 8, **caracterizado** porque las secuencias (iii) son las secuencias que presentan para cualquier sucesión de 100 monómeros contiguos por lo menos 70% de identidad con las secuencias (i) o (ii) respectivamente.
- 45 10. Fragmento nucleotídico que comprende una secuencia nucleotídica que codifica para un polipéptido que presenta, para cualquier sucesión contigua de por lo menos 30 aminoácidos, por lo menos 50% de identidad con una secuencia peptídica seleccionada de entre el grupo constituido por las SEC ID n° 135 y SEC ID n° 137, siendo dicho fragmento nucleotídico diferente de la secuencia HSAC64 del clon humano RG083M05 accesible en la base de datos EMBL, con el número AC000064.
- 50 11. Fragmento nucleotídico según la reivindicación 10, que consiste en una secuencia nucleotídica que codifica para un polipéptido que presenta, para cualquier sucesión contigua de por lo menos 30 aminoácidos, por lo menos 50% de identidad con una secuencia peptídica seleccionada de entre el grupo constituido por las SEC ID n° 135 y SEC ID n° 137.
- 55 12. Fragmento según la reivindicación 11, **caracterizado** porque codifica para un polipéptido que presenta, para cualquier sucesión de por lo menos 30 aminoácidos, por lo menos 70% de identidad con las SEC ID n° 135 y SEC ID n° 137.
- 60 13. Sonda nucleica para la detección de un retrovirus asociado a la esclerosis en placas y/o a la poliartritis reumatoide, **caracterizada** porque es susceptible de hibridarse específicamente sobre cualquier fragmento según cualquiera de las reivindicaciones 7 a 12, que pertenece al genoma de dicho retrovirus.
- 65 14. Sonda según la reivindicación 13, **caracterizada** porque comprende de 10 a 100 nucleótidos.
15. Sonda según la reivindicación 14, **caracterizada** porque comprende de 10 a 30 nucleótidos.

ES 2 331 251 T3

16. Cebador para la amplificación mediante polimerización de un ARN o de un ADN de un retrovirus asociado a la esclerosis en placas y/o a la poliartritis reumatoide, **caracterizado** porque comprende un fragmento según cualquiera de las reivindicaciones 7 a 12.

5 17. Cebador según la reivindicación 16, **caracterizado** porque su secuencia nucleotídica se selecciona de entre las SEC ID n° 127, SEC ID n° 128, SEC ID n° 129, SEC ID n° 132 y SEC ID n° 133.

18. ARN que comprende un fragmento según cualquiera de las reivindicaciones 7 a 12.

10 19. ADN que comprende un fragmento según cualquiera de las reivindicaciones 7 a 12.

20. Vector de replicación y/o de expresión, que comprende un fragmento genómico del material nucleico según cualquiera de las reivindicaciones 1 a 5 o un fragmento según cualquiera de las reivindicaciones 7 a 12.

15 21. Péptido codificado por cualquier marco de lectura abierto que pertenece a un fragmento nucleotídico según cualquiera de las reivindicaciones 7 a 10, que forma o que comprende un determinante antigénico reconocido por unos sueros de pacientes infectados por el virus MSRV-1 y/o en los que el virus MSRV-1 se ha reactivado.

20 22. Péptido según la reivindicación 21, que comprende una secuencia idéntica a una secuencia seleccionada de entre las SEC ID n° 135 y SEC ID n° 137.

25 23. Composición diagnóstica, profiláctica o terapéutica para inhibir la expresión de por lo menos un retrovirus asociado a la esclerosis en placas y/o a la poliartritis reumatoide, que comprende un fragmento nucleotídico según cualquiera de las reivindicaciones 7 a 12.

30 24. Procedimiento para detectar un retrovirus asociado a la esclerosis en placas y/o a la poliartritis reumatoide, en una muestra biológica, **caracterizado** porque se pone en contacto un ARN y/o un ADN que presuntamente pertenece o que procede de dicho retrovirus, o su ARN y/o ADN complementario, con una composición que comprende un fragmento nucleotídico según cualquiera de las reivindicaciones 7 a 12.

35

40

45

50

55

60

65

70

FIG 1

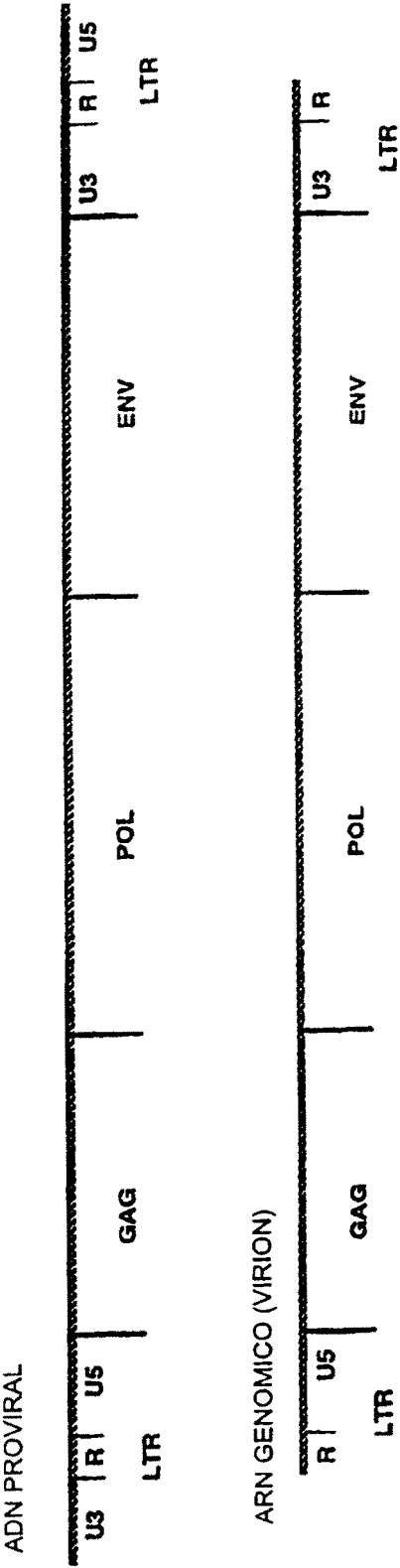


FIG 2

10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	
GCTTATAGAA	GGACCCCTAG	TATGGGGTAA	TCCCCICIGG	GAAACCAAGC	50
A Y R R	T P S	M G .	S P L G	N Q A	
L I E	G P L V	W G N	P L W	E T K P	
L .	K D P .	Y G V I	P S G	K P S	
CCCAGTACTC	AGCAGGAAAA	ATAGAATAGG	AAACCTCACA	AGGACATACT	100
P V L	S R K N	R I G	N L T	R T Y F	
Q Y S	A G K	I E .	E T S Q	G H T	
P S T Q	Q E K .	N R	K P H K	D I L	
TTCCTCCOCT	CCAGATGGCT	AGCCACTGAG	GAAGGAAAAA	TACTTTCACC	150
P P L	Q M A	S H .	G R K N	T F T	
F L P S	R W L	A T E	E G K I	L S P	
S S P	P D G .	P L R	K E K	Y F H L	
TGCAGCTAAC	CAACAGAAAT	TACTTAAAAC	CCTTCACCAA	ACCTTCCACT	200
C S .	P T E I	T . N	P S P N	L P L	
A A N	Q Q K L	L K T	L H Q	T F H L	
Q L T	N R N	Y L K P	F T K	P S T	
TAGGCATTGA	TAGCACCCAT	CAGATGGCOA	AATTATTATT	TACTGGACCA	250
R H .	. H P S	D G Q	I I I	Y W T R	
G I D	S T H	Q M A K	L L F	T G P	
. A L I	A P I	R W P	N Y Y L	L D Q	
GGCCTTTTCA	AAACTATCAA	GAAGATAGTC	AGGGGCTIGT	AAGIGTGCCA	300
P F Q	N Y Q	E D S Q	G L .	S V P	
G L F K	T I K	K I V	R G C E	V C Q	
A F S	K L S R	R .	S G A V	K C A K	
AAGAAATAT					310
K K .					
R N N					
E I					

FIG 2 (continuación)

10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	
CCCTGTATCT	TTAACCTOCT	TGTTAAGTTT	GICTCTTCCA	GAATCAAAC	50
P C I F	N L L	V K F	V S S R	I K T	
P V S	L T S L	L S L	S L P	E S K L	
L Y L	. P P	C . V C	L F Q	N Q N	
TGTAATACTA	CAAAITGTTT	TTCAAATGGA	GCACCAGATG	GAGTCCATGA	100
V K L	Q I V L	Q M E	H Q M	E S M T	
. N Y	K L F	F K W S	T R W	S P .	
C K T T	N C S	S N G	A P D G	V H D	
CTAAGATCCA	CCGTGGACCC	CTGGACCGGC	CTGCTAGCCC	ATGCTCCGAT	150
K I H	R G P	L D R P	A S P	C S D	
L R S T	V D P	W T G	L L A H	A P M	
. D P	P W T P	G P A C	. P M L R C		
GTTAATGACA	TTGAAGGCAC	CCCTCCCGAG	GAAATCTCAA	CTGCACAACC	200
V N D I	E G T	P P E	E I S T	A Q P	
L M T	L K A P	L P R	K S Q	L H N P	
. . H	. R H	P S R G	N L N	C T T	
OCTACTATGC	CCCAATTCAG	CGGGAACCAG	TTAGAGCGGT	CATCAGCCAA	250
L L C	P N S A	G S S	. S G	H Q P T	
Y Y A	P I Q	R E A V	R A V	I S Q	
P T M P	Q F S	G K Q	L E R S	S A N	
CCTCCCCAAC	AGCACTTGGG	TTTTCCIGTT	GAGAGGGGGG	ACTGAGAGAC	300
S P T	A L G	F S C	. E G G	L R D	
P P Q Q	H L G	F P V	E R G D	. E T	
L P N	S T W V	F L L	R G G	T E R Q	
AGGACTAGCT	GGATTTCCTA	GGCCAAAGAA	GAATCCCTAA	GOCTAGCTGG	350
R T S W	I S .	A N E	E S L S	L A G	
G L A	G F P R	P T K	N P .	A . L G	
D . L	D F L	G Q R R	I P K	P S W	

FIG 3

10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	
GAAGGTGACT	GCATCCACT	CTAAACATGG	GGCTTGCAAC	TTAGCTCACA	400
K V T	A S T S	K H G	A C N	L A H T	
R .	L H P P	L N M G	L A T	. L T	
E G D C	I H L	. T W	G L Q L	S S H	
CCCGACCAAT	CAGAGAGCTC	ACTAAAATGC	TAATTAGGCA	AAAATAGGAG	450
R P I	R E L	T K M L	I R Q	K . E	
P D Q S	E S S	L K C	. L G K	N R R	
P T N	Q R A H	. N A N	. A	K I G G	
GTAAGAAT	AGCCAATCAT	CTATTGCCCTG	AGAGCACAGC	GGGAGGGACA	500
V K K .	P I I	Y C L	R A Q R	E G Q	
. R N	S Q S S	I A .	E H S	G R D K	
K E I	A N H	L L P E	S T A	G G T	
AGGATCGGGA	TATAAATCCA	GGCATTGAG	COGGCAACGG	CAACCCCTT	550
G S G	Y K P R	H S S	R Q R	Q P P L	
D R D	I N P	G I R A	G N G	N P L	
R I G I	. T Q	A F E	P A T A	T P F	
TGGGTCCCT	CCCTTGTAT	GGCGCTCTG	TTTCACTCT	ATTCACTCT	600
G P L	P L Y	G R S V	F T L	F H S	
W V P S	L C M	G A L	F S L Y	F T L	
G S P	P F V W	A L C	F H S	I S L Y	
ATTAATCTT	GCAACTGAAA	AAAAAAAAAA	AAAAA		635
I K S C N	. K	K K K K	K		
L N L	A T E K	K K K K	K		
. I L	Q L K	K K K K	K		

FIG 4

10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	
ATGGCCCTCC	CTTATCATA	TTTTCTCTT	ACTGTTCTCT	TACCCCTTT	50
M A L P	Y H T	F L F	T V L L	P P F	
W P S	L I I L	F S L	L F S	Y P L S	
G P P	L S Y	F S L Y	C S L	T P F	
CGCTCTCACT	GCACCCCTC	CATGCTGCTG	TACAACCACT	AGCTCCCTT	100
A L T	A P P P	C C C	T T S	S S P Y	
L S L	H P L	H A A V	Q P V	A P L	
R S H C	T P S	M L L	Y N Q	. L P L	
ACCAAGAGTT	TCTATGAAGA	ACGCGGCTC	CTGGAATAT	TGATGCCCCA	150
Q E F L	. R	T R L P	G N I	D A P	
T K S F	Y E E	R G F	L E I L	M P H	
P R V	S M K N	A A S	W K Y	. C P I	
TCATATAGGA	GTTTATCTAA	GGGAACTCC	ACCTTCACTG	OCCACAACCA	200
S Y R S	L S K	G N S	T F T A	H T H	
H I G	V Y L R	E T P	P S L	P T P I	
I . E	F I .	G K L H	L H C	P H P	
TATGCCCCGC	AACTGCTATA	ACTCTGCCAC	TCTTTGCAATG	CATGCAATA	250
M P R	N C Y N	S A T	L C M	H A N T	
C P A	T A I	T L P L	F A C	M Q I	
Y A P Q	L L .	L C H	S L H A	C K Y	
CTCATIATIG	GACAGGAAA	ATGATTAATC	CTAGTTGTCC	TGGAGGACTT	300
H Y W	T G K	M I N P	S C P	G G L	
L I I G	Q G K	. L I	L V V L	E D L	
S L L	D R E N D	. S .	L S	W R T W	
GGAGCCACTG	TCTGTTGGAC	TACTTCACC	CATACCAGTA	TGCTCGATGG	350
G A T V	C W T	Y F T	H T S M	S D G	
E P L	S V G L	T S P	I P V	C L M G	
S H C	L L D	L L H P	Y Q Y	V . W	

FIG 4 (continuación)

10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	
GGGTGGAATT	CAAGGTCAGG	CAAGAGAAAA	ACAAGTAAAG	GAAGCAATCT	400
G G I Q G Q A	R E K Q V K	E A I S			
V E F K V R	Q E K N	K * R K Q S			
G W N S	R S G K R K	T S K G	S N L		
CCCAACTGAC	CCGGGGACAT	AGCACCCCTA	GCCCCTACAA	AGGACTAGTT	450
Q L T R G H	S T P S	P Y K G L V			
P N * P G D I	A P L A P T K	D * F			
P T D P G T *	H P * P L Q	R T S S			
CTCTCAAAAC	TACATGAAAC	CCTCCGTACC	CATACTCGCC	TGGTGAGCCT	500
L S K L H E T	L R T H T R L	V S L			
S Q N Y M K P	S V P I L A	W * A Y			
L K T T * N	P P Y P Y S P	G E P			
ATTTAATACC	ACCCTCACTC	GGCTCCATGA	GGTCTCAGCC	CAAAACCCTA	550
F N T T L T R	L H E V S A	Q N P T			
L I P P S L	G S M R S Q P	K T L			
I * Y H P H S	A P * G L S P	K P Y			
CTAACTGTTG	GATGTGCCTC	CCCCTGCACT	TCAGGCCATA	CATTTCAATC	600
N C W M C L	P L H F R P Y	I S I			
L T V G C A S	P C T S G H T	F Q S			
* L L D V P P	P A L Q A I	H F N P			
CCTGTTCTTG	AACAATGGAA	CAACTTCAGC	ACAGAAATAA	ACACCACTTC	650
P V P E Q W N	N F S T E I N	T T S			
L F L N N G T	T S A Q K *	T P L P			
C S * T M E	Q L Q H R N K	H H F			
CGTTTTAGTA	GGACCTCTTG	TTTCCAATCT	GGAAATAACC	CATACCTCAA	700
V L V G P L	V S N L E I T	H T S N			
F * * D L L	F P I W K *	P I P Q			
R F S R T S C	F Q S G N N P	Y L K			

FIG 4 (continuación)

10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	
ACCTCACCIG	TGTAATAATT	AGCAATACTA	TAGACACAAC	CAGCTCCCAA	750
L T C	V K F	S N T	I D T	T S S Q	
T S P V	. N L	A I L	. T Q P	A P N	
P H L	C K I	. Q Y Y	R H N	Q L P M	
TGCATCAGGT	GGGTAACAAC	TCCACAAGA	ATAGTCTGCC	TACCCCTCAGG	800
C I R W	V T P	P T R	I V C L	P S G	
A S G G	. H L	P H E	. S A	Y P Q E	
H Q V	G N T	S H T N	S L P	T L R	
AATATTTTTT	GTCIGGGTA	CCTCAGCCTA	TCATGTITIG	AATGGCTCTT	850
I F F	V C G T	S A Y	H C L	N G S S	
Y F L	S V V	P Q P I	I V .	M A L	
N I F C	L W Y	L S L	S L F E	W L F	
CAGAATCTAT	GIGCTTCCIC	TCATICTIAG	TGCCCCCTAT	GACCATCTAC	900
E S M	C F L	S F L V	P P M	T I Y	
Q N L C	A S S	H S .	C P L .	P S T	
R I Y	V L P L	I L S	A P Y	D H L H	
ACTGAACAAG	ATTATACAA	TCATGTGTA	OCTAAGCCCC	ACAACAAAAG	950
T E Q D	L Y N	H V V	P K P H	N K R	
L N K	I Y T I	M S Y	L S P	T T K E	
. T R	F I Q	S C R T	. A P	Q Q K	
AGTACCCATT	CTTCCTTTTG	TTATCAGAGC	AGGAGTGCTA	GGCAGACTAG	1000
V P I	L P F V	I R A	G V L	G R L G	
Y P F	F L L	L S E Q	E C .	A D .	
S T H S	S F C	Y Q S	R S A R	Q T R	
GTACTGGCAT	TGGCAGTATC	ACAACCICIA	CTCAGTICTA	CTACAAACTA	1050
T G I	G S I	T T S T	Q F Y	Y K L	
V L A L	A V S	Q P L	L S S T	T N Y	
Y W H	W Q Y H	N L Y	S V L	L Q T I	

FIG 4 (continuación)

10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	
TCTCAAGAAA	TAAATGGTGA	CATCGAACAG	GTCACTGACT	COCTGGTCAC	1100
S Q E I	N G D M	E Q V T	D S L V	T	
L K K .	M V T W	N R S L	T P W S	P	
S R N K	W . H	G T G H .	L P G H		
CTTGCAAGAT	CAACTTAACT	COCTAGCAGC	AGPAGTCCTT	CAAAATCGAA	1150
L Q D Q	L N S L	A A V V	L Q N R	R	
C K I N	L T P .	Q Q . S	F K I E		
L A R S	T . L	P S S S	S P S K	S K	
GAGCTTTAGA	CITGCTAACC	GCCAAAAGAG	GGGGAACCTG	TTTATTTTITA	1200
A L D L	L T A K	R G G T	C L F L		
E L . T	C . P P	K E G E	P V Y F .		
S F R L	A N R Q	K R G N	L F I F	R	
GGAGAAGAAC	GCTGTIATTA	TGTIATICAA	TCCAGATTG	TCACTGAGAA	1250
G E E R	C Y Y V	N Q S R	I V T E	K	
E K N A	V I M L	I N P E	L S L R	K	
R R T L	L L C .	S I Q N	C H . E		
AGTTAAAGAA	ATTGAGATC	GAATACAATG	TAGACCAGAG	GAGCTTCAAA	1300
V K E I	R D R I	Q C R A	E E L Q	N	
L K K F	E I E Y	N V E Q	R S F K		
S . R N	S R S N	T M . S	R G A S	K	
ACACOGAAG	CTGGGGCCTC	CTCAGCCAAT	GGATGCCCTG	GGTTCICCCC	1350
T E R W	G L L S	Q W M P	W V L P		
T P N A	G A S S	A N G C	P G F S	P	
H R T L	G P P Q	P M D A	L G S P	L	
TTCTTAGGAC	CTCTAGCAGC	TCTAATATIG	TTACTCCTCT	TIGGACCCCTG	1400
F L G P	L A A L	I L L L	L L F G	P C	
S . D L .	Q L . Y	C Y S S	L D P V		
L R T S	S S S S	N I V T	P L W T	L	

FIG 4 (continuación)

10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	
TATCITTAAC	CTCCTTGTTA	AGTTIGICTC	TCCAGAATT	GAAGCTGTAA	1450
I F N	L L V K	F V S	S R I	E A V K	
S L T	S L L	S L S L	P E L	K L .	
Y L .	P P C .	V C L	F Q N .	S C K	
AGCTACAGAT GGTCTTACAA ATGGAACCCC A					1481
L Q M	V L Q	M E P			
S Y R W	S Y K	W N P			
A T D	G L T N	G T P			

FIG 5

	10	20	30	40	50	
	1234567890	1234567890	1234567890	1234567890	1234567890	
	TCAAATCGA	AGAGCTTAG	ACTTGCTAAC	CGCCAAAAGA	GGGGAACCT	50
	S K S K	S F R	L A N	R Q K R	G N L	
	Q N R	R A L D	L L T	A K R	G G T C	
	K I E E	L T C	P P	K E	G E P	
	GTTTATTTT	AGGGGAAGAA	TGCTGTAGT	ATGTTAATCA	ATCTGGAATC	100
	F I F	R G R M	L L V C	S I W N H		
	L F L	G E E C C	Y V N Q	S G I		
	V Y F	G K N A V S	M L I N	L E S		
	ATTACTGAGA	AAGTTAAAGA	AATTTGAGAT	CGAATATAAT	GTAGAGCAGA	150
	Y . E S .	R N L R S	N I M .	S R		
	I T E K	V K E I .	D R I .	C R A E		
	L L R	K L K K	F E I E	Y N V E Q R		
	GGACCTTCAA	AACACTGCAC	CCTGGGGCCT	CCTCAGCCAA	TGGATGCCCT	200
	G P S K	H C T	L G P	P Q P M	D A L	
	D L Q	N T A P	W G L	L S Q	W M P W	
	T F K	T L H	P G A S	S A N	G C P	
	GGACTCTCC	CTTCTTAGGA	CCTCTAGCAG	CTATAATATT	TTTACTCCTC	250
	D S P	L L R T	S S S	Y N I	F T P L	
	T L P	F L G	P L A A	I I F	L L L	
	G L S P	S . D L .	Q L .	Y F	Y S S	
	TTTGGACCT	GTACTTCAA	CTTCTTGT	AAGTTTGTCT	CTTCCAGAAT	300
	W T L	Y L Q	L P C .	V C L	F Q N	
	F G P C	I F N	F L V	K F V S	S R I	
	L D P	V S S T	S L L	S L S	L P E L	
	TGAAGCTGTA	AAGCTACAAA	TAGTCTTCA	AATGGAACCC	CAGATGCAGT	350
	. S C K	A T N	S S S	N G T P	D A V	
	E A V	K L Q I	V L Q	M E P	Q M Q S	
	K L .	S Y K .	F F K	W N P	R C S	

FIG 5 (continuación)

10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	
CCATGACTAA	AATCTACCGT	GGACCCCTGG	ACCGGCTGTC	TAGACTATGC	400
H D .	N L P W	T P G	P A C	. T M L	
M T K	I Y R	G P L D	R P A	R L C	
P . L K	S T V	D P W	T G L L	D Y A	
TCTGATGTTA	ATGACATTGA	AGTCACCCCT	CCCGAGGAAA	TCTCAACTGC	450
. C .	. H .	S H P S	R G N	L N C	
S D V N	D I E	V T P	P E E I	S T A	
L M L	M T L K	S P L	P R K	S Q L H	
ACAACCCCTA	CTACACTCCA	ATTCACTAGG	AAGCAGTTAG	AGCAGTTGTC	500
T T P T	T L Q	F S R	K Q L E	Q L S	
Q P L	L H S N	S V G	S S .	S S C Q	
N P Y	Y T P	I Q .	E A V R	A V V	
AGCCAACCTC	CCCAACAGTA	CTTGGGTTTT	CCTGTGAGAG	GGTGGACTGC	550
A N L	P N S T	W V F	L L R	G W T E	
P T S	P T V	L G F S	C .	E G G L	
S Q P P	Q Q Y	L G F	P V E R	V D .	
AGAGACAGGA	CTAGCTGGAT	TTCCTAGGCT	GACTAAGAAT	CCCAAGCCT	600
R Q D .	L D F L G .	L R I	P K P		
R D R T	S W I S .	A D .	E S X S L		
E T G	L A G F	P R L	T K N	P X A X	
ANCTGGGAAG	GTGACCGCAT	CCATCTTTTAA	ACATGGGGCT	TGCAACTTAG	650
X W E G	D R I H L .	T W G L	Q L S		
X G K	V T A S	I F K	H G A	C N L A	
L G R .	P H P S L N	M G L	A T .		
CTCACACCCG	ACCAATCAGA	GAGCTCACTA	AAATGCTAAT	CAGGCAAAAA	700
S H P	T N Q R	A H .	N A N	Q A K T	
H T R	P I R	E L T K	M L I	R Q K	
L T P D	Q S E	S S L	K C .	S G K N	

FIG 5 (continuación)

10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	
CAGGAGGTAA	AGCAATAGCC	AATCATCTAT	TGCCTGAGAG	CACAGCGGGA	750
G G K	A I A	N H L L	P E S	T A G	
Q E V K	Q . P	I I Y	C L R A	Q R E	
R R .	S N S Q	S S I	A . E	H S G K	
AGGACAAGGA	TTGGGATATA	AACTCAGCCA	TTCAAGCCAG	CAACAGCAAC	800
R T R I	G I .	T Q A	F K P A	T A T	
G Q G	L G Y K	L R H	S S Q	Q Q Q P	
D K D	W D I	N S G I	Q A S	N S N	
CCCCTTGGG	TCCCCIOCCA	TIGTIATGGGA	GCTCIGITTT	CACTCIATTT	850
P F G	S P P I	V W E	L C F	H S I S	
P L G	P L P	L Y G S	S V F	T L F	
P L W V	P S H	C M G	A L F S	L Y F	
CACTCIATTA	AATCATGCAA	CTCCACTICTT	CTGGTCCGIG	TTTTTIATGG	900
L Y .	I M Q	L H S S	G P C	F L W	
H S I K	S C N	C T L	L V R V	F Y G	
T L L	N H A T	A L F	W S V	F F M A	
CTCAAGCTGA	GCTTTTGTTC	GCCATCCACC	ACTGCTGITT	GCCACCGTCA	950
L K L S	F C S	P S T	T A V C	H R H	
S S .	A F V R	H P P	L L F	A T V T	
Q A E	L L F	A I H H	C C L	P P S	
CAGACCCGCT	GCTIGACTTCC	ATCCCTTTGG	ATCCAGCAGA	GTTGTCCTIG	1000
R P A	A D F H	P F G	S S R	V S T V	
D P L	L T S	I P L D	P A E	C P L	
Q T R C	. L P	S L W	I Q Q S	V H C	
TGCTCCTGAT	CCAGCGAGGT	ACCCATIGCC	ACTCCCGATC	AGGCTAAAGG	1050
L L I	Q R G	T H C H	S R S	G . R	
C S .	S S E V	P I A	T P D Q	A K G	
A P D	P A R Y	P L P	L P I	R L K A	

FIG 5 (continuación)

10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	
CTTGCCATIG	TTCCTGCATG	GCTAAGTGCC	TGGGTTTGIC	CTAATAGAAC	1100
L A I V	P A W	L S A	W V C P	N R T	
L P L	F L H G	. V P	G F V	L I E L	
C H C	S C M	A K C L	G L S	. . N	
TGAACACTGG	TCACTGGGTT	CCATGGTICT	CTTCCATGAC	CCACGGCTTC	1150
E H W	S L G S	M V L	F H D	P R L L	
N T G	H W V	P W F S	S M T	H G F	
. T L V	T G F	H G S	L P .	P T A S	
TAATAGAGCT	ATAACACTCA	CCGCATGGCC	CAAGATTCCA	TTCCTTGGTA	1200
I E L	. H S	P H G P	R F H	S L V	
. . S Y	N T H	R M A	Q D S I	P W Y	
N R A	I T L T	A W P	K I P	F L G I	
TCTGTGAGGC	CAAGAACCCC	AGGTCAGAGA	ANGTGAGGCT	TGCCAOCATT	1250
S V R P	R T P	G Q R	X . G L	P P F	
L . G	Q E P Q	V R E	X E A	C H H L	
C E A	K N P	R S E X	V R L	A T I	
TGGGAAGTGG	CCCCTGCOCA	TTTTGGTAGC	GGCCCCACCAC	CATCTTGGGA	1300
G K W	P T A I	L V A	A H H	H L G S	
G S G	P L P	F W .	R P T T	I L G	
W E V A	H C H	F G S	G P P P	S W E	
GCTGTGGGAG	CAAGGATCCC	CCAGTAACA			1329
C G S	K D P	P V T			
A V G A	R I P	Q .			
L W E	Q G S P	S N			

FIG 6

10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	
CCTAGAACGT	ATTCTGGAGA	ATTGGACCA	ATGTGACACT	CAGACCTAA	50
P R T Y	S G E	L G P	M . H S	D A K	
L E R	I L E N	W D Q	C D T	Q T L R	
. N V	F W R	I G T N	V T L	R R .	
GAAAGAAAG	ATTTATATTC	TTCTGCACTA	CCGCTGCC	ACAATATCT	100
K E T	I Y I L	L Q Y	R L A	T I S S	
K K R	F I F	F C S T	A W P	Q Y P	
E R N D	L Y S	S A V	P P G H	N I L	
CTTCAAGGA	GAGAACTTG	GCTTCTGAG	GGAAGTATAA	ATTATAACAT	150
S R E	R N L	A S .	G K Y	K L . H	
L Q G	R E T W	L P E	G S I N	Y N I	
F K G	E K P G	F L R	E V .	I I T S	
CATCTACAG	CTAGACTCT	TCTGTAGAAA	GGAGGCCAAA	TGAGTGAAG	200
H L T	A R P L	L . K	G G Q M	E . S	
I L Q	L D L F	C R K	E G K	W S E V	
S Y S	. T S	S V E R	R A N	G V K	
TGCCATATGT	GCAAACTTTC	TTTTCATTA	GAGACAATC	ACAATTATGT	250
A I C	A N F L	F I K	R Q L	T I M .	
P Y V	Q T F	F S L R	D N S	Q L C	
C H M C	K L S	F H .	E T T H	N Y V	
AAAAAGTGTG	GTTTATGCC	TACAGGAGC	CCTCAGAGTC	CACCTCCCTA	300
K V W	F M P	Y R K P	S E S	T S L	
K K C	G L C P	T G S	P Q S P	P P Y	
K S V	V Y A L	Q E A	L R V	H L P T	
CCCCAGGTC	CCCTCCCGA	CTCCTTCTC	AACTAATAAG	GACCCCCCT	350
P Q R P	L P D	S F L	N . .	G P P F	
P S V	P S P T	P S S	T N K	D P P L	
P A S	P P R	L L P Q	L I R	T P L	
TAAACAAAC	GGTCCAAAG	GAGATGACA	AGGGGTAAA	CAATGAAACA	400
N P N	G P K G	D R Q	R G K	Q . T K	
T Q T	V Q K	E I D K	G V N	N E P	
. P K R	S K R	R . T	K G .	T M N Q	
AAGATGCCA	ATATTCGGG	ATTATGCCC	CTCCAGGAG	TGAGAGGAG	450
E C Q	Y S P	I M P P	P S S	E R R	
K S A N	I P R	L C P	L Q A V	R G G	
R V P	I F P D	Y A P	S K Q .	E E E	
AGAATTCGC	CCAGCCAGG	TGCCGTACC	TTTTTCTCTC	TCAGACTTAA	500
R I R P	S Q S	A C T	F F S L	R L K	
E F G	P A R V	P V P	F S L	S D L K	
N S A	Q P E	C L Y L	F L S	Q T .	

ES 2 331 251 T3

FIG 6 (continuación)

	10	20	30	40	50	
	1234567890	1234567890	1234567890	1234567890	1234567890	
	AGCAAATTAA	AATAGACCTA	GGTAAATICT	CAGATAAACC	TGACCGCTAT	550
	A N .	N R P R .	I L R .	P .	R L Y	
	Q I K	I D L	G K F S	D N P	D G Y	
	S K L K .	T .	V N S	Q I T L	T A I	
	ATTGATGTTT	TACAAGGGTT	AGGACAATCC	TTTGATCTGA	CATGGAGAGA	600
	. C F	T R V	R T I L .	S D	M E R	
	I D V L	Q G L	G Q S	F D L T	W R D	
	L M F	Y K G .	D N P	L I .	H G E I	
	TATAATGTTA	CTACTAAATC	AGACACTAAC	CCCAAATGAG	AGAAGTGOOG	650
	Y N V T	T K S	D T N	P K .	E K C R	
	I M L	L L N Q	T L T	P N E	R S A A	
	. C Y Y .	I R H .	P	Q M R	E V P	
	CTGTAAGTGC	AACCGAGAG	TTTGGGATC	TTTGGTACT	CAGTCAGGCC	700
	C N C	S P R V	W R S	L V S	Q S G Q	
	V T A	A R E	F G D L	W Y L	S Q A	
	L .	L Q	P E S	L A I	F G I S .	V R P
	AACAATGCA	TGACAACAGA	GGAAAGAACA	ACTCCACAG	GCCAGCAGGC	750
	Q .	D D N R	G K N N	S H R	P A G	
	N N R M	T T E	E R T	T P T G	Q Q A	
	T I G .	Q Q R	K E Q	L P Q	A S R Q	
	AGTTCOCAGT	GTAGACCCTC	ATTGGGACAC	AGAATCAGAA	CATGGAGATT	800
	S S Q C	R P S	L G H	R I R T	W R L	
	V P S	V D P H	W D T	E S E	H G D W	
	F P V .	T L	I G T Q	N Q N	M E I	
	GGTGGCACAA	ACATTTGCTA	ACTTGGTGC	TAGAAGGACT	GAGGAAACT	850
	V P Q	T F A N	L R A	R R T E E N .		
	C H K	H L L	T C V L	E G L	R K T	
	G A T N	I C .	L A C .	K D .	G X L	
	AGGAGAAGC	CTATGAATTA	CTCAATGATG	TCCACTATA	CACAGGGAAA	900
	E E A	Y E L	L N D V	H Y N	T G K	
	R K K P	M N Y	S M M	S T I T	Q G K	
	G R S L .	I T Q .	C P L .	H R E R		
	GGAGAAAAT	CTTACTGCTT	TTCTGGACAG	ACTAAGGGAG	GCATTGAGGA	950
	G R K S	Y C F	S G Q	T K G G	I E E	
	E E N	L T A F	L D R	L R E	A L R K	
	K K I	L L L	F W T D .	G R H .	G	
	AGCATACCTC	CCTGTCACT	GACTCTATG	AAGGCCAACT	AATCTTAAAG	1000
	A Y L	P V T .	L Y .	R P T	N L K G	
	H T S	L S P	D S I E	G Q L	I L K	
	S I P P	C H L	T L L	K A N .	S . R	

FIG 6 (continuación)

10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	
GATAAGTTTA	TCACTCAGTC	ACCTGCAGAC	ATTAGAAAA	ACTTCAAAG	1050
. V Y H S V	S C R H	. K K	L Q K		
D K F I T Q S	A A D	I R K N	F K S		
I S L S L S Q	L Q T	L E K	T S K V		
TCTGCTTAG	GCCCCGAGCA	GAAGTTAGAA	ACCCTATTTA	ACTTGGCATT	1100
S A L G P E Q	N L E	T L F N	L A S		
L P . A R S R T	. K P Y L	T W H P			
C L R P G A	E L R N	P I .	L G I		
CTCAGTTTTT	TATAATAGAG	ATCAGGAGCA	CCAGCCGAAA	CGGGACAAC	1150
S V F Y N R D	Q E E	Q A K	R D K R		
Q F F I I E	I R R S	R R N	G T N		
L S F L . .	R S G G	A G E T	G Q T		
GGATAAAAA	AAAAAGGGG	GGTCCACTAC	TTTAGTCATG	GCCCTCAGGC	1200
D K K K R G	G P L L	. S W	P S G		
G I K K K G G	V H Y	F S H G	P Q A		
G . K K K G G	S T T	L V M	A L R Q		
AAGCAGACTT	TGAGGCTCT	GCAAAAGGA	AAAGCTGGCC	AAATCAAATG	1250
K Q T L E A L	Q K G	K A G Q	I K C		
S R L W R L C	K R E	K L G	K S N A		
A D F G G S	A K G K	S W A	N Q M		
CCATATAGGG	CTGGCTTCCA	GTCGGGCTA	CAAGGACTT	TAAAAAGA	1300
L I G L A S S	A V Y	K D T	L K K I		
. . G W L P	V R S T	R T L	. K R		
P N R A G F Q	C G L	Q G H F	K K D		
TTATCCAGT	AGAAATAGC	CGCCCCCTG	TCCATGCCCC	TTAGGTCAAG	1350
I Q V E I S	R P L V	H A P	Y V K		
L S K . K . A	A P L	S M P L	T S R		
Y P S R N K P	P P C	P C P	L R Q G		
GGATCAGCTG	GAAGCCCCAC	TGCCCCAGGG	GATGAGATA	CTCGAGTCA	1400
G I T G R P T	A P G	D E D T	L S Q		
E S L E G P L	P Q G	M K I	L . V R		
N H W K A H	C P R G	. R Y	S E S		
GAAGCATTG	ACCAGATGAT	CCAGCAGCAG	GACTGAGGT	CCCCGGGGG	1450
K P L T R . S	S S R	T E G	A R G E		
S H . P D D	P A A G	L R V	P G A		
E A I N Q M I	Q Q Q	D . G C	P G R		
AGCCCCAGCC	CATGOCATCA	CCCTCAGAGA	GGCCCCGGTA	TGTTTGACCA	1500
R Q P M P S	P S Q S	P G Y	V . P		
S A S P C H H	P H R	A P G M	F D H		
A P A H A I T	L T E	P R V	C L T I		

FIG 6 (continuación)

10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	
TTGAGAGCCA	A				1511
L	R	A			
.	E	P			
E	S	Q			

ES 2 331 251 T3

FIG 7

	10	20	30	40	50	
	1234567890	1234567890	1234567890	1234567890	1234567890	
	ATGGGCAGCA	GCCATCATCA	TCATCATCAC	ACCAGCGGCC	TGGTGGCGCG	50
	M G S S	H H H	H H H	S S G L	V P R	
	CGGCAGCCAT	ATGGCTAGCA	TGACTGGTGG	ACAGCAAATG	CGTGGATCC	100
	G S H	M A S M	T G G	Q Q M	G R I L	
	TAGAAGTAT	TCTGGAGAAT	TGGGACCAAT	GTGACACTCA	GACGCTAAGA	150
	E R I	L E N	W D Q	C D T	Q T L R	
	AAGAAACGAT	TTATATTCTT	CTGCAGTACC	GCCTGGCCAC	AATATCCTCT	200
	K K R F	I F F	C S T	A W P Q	Y P L	
	TCAAGGAGA	GAAACCTGCG	TTCTGAGGG	AAGTATAAAT	TATAACATCA	250
	Q G R	E T W L	P E G	S I N	Y N I I	
	TCTTACAGCT	AGACCTCTTC	TGTAGAAAGG	AGGGCAAATG	GAGTGAAGTG	300
	L Q L	D L F	C R K E	G K W	S E V	
	CCATATGTGC	AAACTTTCIT	TTCAATTAAGA	GACAACTCAC	AATATGTAA	350
	P Y V Q	T F F	S L R	D N S Q	L C K	
	AAAGTGTGGT	TTATGCCCTA	CAGGAAGGCC	TCAGAGTCCA	CCTUCCTACC	400
	K C G	L C P T	G S P	Q S P	P P Y P	
	CCAGGTGCC	CTCCCCGACT	CCTTCTCAA	CTAATTAAGGA	CCCCCTTTA	450
	S V P	S P T	P S S T	N K D	P P L	
	ACCCAAACGG	TCCAAAAGGA	GATAGACAAA	GGGGTAAACA	ATGAAOCAA	500
	T Q T V	Q K E	I D K	G V N N	E P K	
	GAGTGGCAAT	ATTCOOOGAT	TATGCCCCCT	CCAAGCAGTG	AGAGGAGGAG	550
	S A N	I P R L	C P L	Q A V	R G G E	
	AATTCGGGCC	AGCCAGAGTG	CCTGTACCTT	TTTCTCTCTC	AGACTTAAG	600
	F G P	A R V	P V P F	S L S	D L K	
	CAAATTAAAA	TAGACCTAGG	TAAATTCICA	GATACCCCTG	ACGGCTATAT	650
	Q I K I	D L G	K P S	D N P D	G Y I	
	TGATGTTTTA	CAAGGGTAG	GACAATCCTT	TGATCTGACA	TGGAGAGATA	700
	D V L	Q G L G	Q S F	D L T	W R D I	
	TAATGTACT	ACTAAATCAG	ACACTAACCC	CAATGAGAG	AAGTGGCGCT	750
	M L L	L N Q	T L T P	N E R	S A A	
	GTAACTCAG	COGAGAGTT	TGGGATCCTT	TGGTATCTCA	GTCAGGCCAA	800
	V T A A	R E F	G D L	W Y L S	Q A N	

FIG 7 (continuación)

10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	
CATAGGATG	ACAACAGAGG	AAAGAACAAC	TCCACAGGC	CAGCAGGCAG	850
N R M	T T E E	R T T	P T G	Q Q A V	
TTCOCAGTGT	AGACCCCTCAT	TGGGACACAG	AATCAGAACA	TGGAGATTGG	900
P S V	D P H	W D T E	S E H	G D W	
TGCCACAAAC	ATTTGCTAAC	TTGGTGCTA	GAAGGACTGA	GGAAAAGTAG	950
C H K H	L L T	C V L	E G L R	K T R	
GAGAAGCCT	ATGAATTACT	CAATGATGTC	CACTATAACA	CAGGGAAGG	1000
K K P	M N Y S	M M S	T I T	Q G K E	
AAGAAAATCT	TACTGCTTTT	CTGGACAGAC	TAAGGGAGGC	ATTCAGGAAG	1050
E N L	T A F	L D R L	R E A	L R K	
CATAOCTOCC	TGTCACCTGA	CTCTATTGAA	GCCCACTAA	TCTTAAAGGA	1100
H T S L	S P D	S I E	G Q L I	L K D	
TAAGTTTATC	ACTCAGTCAG	CTGCAGACAT	TAGAAAAAC	TTCAAAGTC	1150
K F I	T Q S A	A D I	R K N	F K S L	
TGCTAAGCT	TGGGGCCGCA	CTGCAGCAC	ACCAACCACCA	CCACTGAGAT	1200
P K L	A A A	L E H H	H H H	H . D	
CCGGCTGCTA	ACAAAGCCCG	AAAGGAGCT	GAGTTGGCTN	GTGGCNA	1247
P A A N	K A R	K E A	E L A X	G	

ES 2 331 251 T3

FIG 8

10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	
ATGGCTAGCA	TGACTGGTGG	ACAGCAAATG	GGTGGATCC	TAGAAGGTAT	50
M A S M	T G G	Q Q M	G R I L	E R I	
TCTGGAGAAT	TGGACCAAT	GTGACACTCA	GAGCCTAAGA	AAGAAAGGAT	100
L E N W	D Q C	D T Q	T L R	K K R F	
TTATATTCTT	CTGCAGTACC	GCCTGGCCAC	AATATCTCT	TCAAGGGAGA	150
I F F C	S T A	W P Q	Y P L	Q G R	
GAAACTGGC	TTCCTGAGGG	AAGTATAAAT	TATAACATCA	TCTTACAGCT	200
E T W L	P E G	S I N	Y N I	I L Q L	
AGACTCTTC	TGTAGAAAGG	AGGGCAAATG	GAGTGAAGTG	CCATATGTCC	250
D L F C	R K E	G K W	S E V	P Y V Q	
AACTTTCTT	TTCATTAAGA	GACAATCAC	AATTATGTAA	AAAGTGTGGT	300
T F F S	L R D	N S Q	L C K	K C G	
TTATGGCCTA	CAGGAAGCC	TCAGAGTCCA	CCTCCCTACC	CCAGGGTCCC	350
L C P T	G S P	Q S P	P P Y P	S V P	
CTCCCGACT	CCTTCCTCAA	CTAATAAGGA	CCCCCTTTA	ACCCAAACGG	400
S P T P	S S T	N K D	P P L	T Q T V	
TOCAAAGGA	GATAGACAAA	GGGTAACA	ATGAACAAA	GAGTGGCAAT	450
Q K E I	D K G	V N N	E P K	S A N	
ATTCCCGAT	TATGGCCCT	CCAAGCAGTG	AGAGGAGGAG	AATTCGGCC	500
I P R L	C P L	Q A V	R G G E	F G P	
AGCCAGAGTG	CCTGTAOCTT	TTTCTCTCTC	AGACTTAAAG	CAAATTA AAA	550
A R V P	V P F	S L S	D L K	Q I K I	
TAGACCTAGG	TAAATTTCTA	GATTAOCTG	ACGGCTATAT	TGATGTTTTA	600
D L G K	F S D	N P D	G Y I	D V L	
CAAGGTTAG	GACAAOCTT	TGATCTGACA	TGGAGAGATA	TATIGTFACT	650
Q G L G	Q S F	D L T	W R D	I M L L	
ACTAAATCAG	ACACTAAOCC	CAAATGAGAG	AAGTGGCGCT	GTAACITGCG	700
L N Q T	L T P	N E R	S A A	V T A A	
COOGAGAGTT	TGGGATCTT	TGGTATCTCA	GTCAGGCCAA	CAATAGGATG	750
R E F G	D L W	Y L S	Q A N	N R M	
ACAACAGAGG	AAGAACAAC	TOCCACAGGC	CAGCAGGCAG	TTOCCAGTGT	800
T T E E	R T T	P T G	Q Q A	V P S V	

FIG 8 (continuación)

10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	
AGACCOCTCAT	TGGGACACAG	AATCAGAACA	TGGAGATTGG	TGCCACAAAC	850
D P H	W D T E	S E H	G D W	C H K H	
ATTTGCTAAC	TTGGTIGCTA	GAAGGACTGA	GGAAACTAG	GAAGAAGCCT	900
L L T	C V L	E G L R	K T R	K K P	
ATGAATTACT	CAATGATGTC	CACTATAACA	CAGGAAAGG	AAGAAAATCT	950
M N Y S	M M S	T I T	Q G K E	E N L	
TACTGCTTTT	CTGGACAGAC	TAAGGGAGGC	ATIGAGGAAG	CATACTCCC	1000
T A F	L D R L	R E A	L R K	H T S L	
TGTCAOCTGA	CICTATTGAA	GGCCAATAA	TCTTAAAGGA	TAAGTTTATC	1050
S P D	S I E	G Q L I	L K D	K F I	
ACTCAGTCAG	CTGCAGACAT	TAGAAAAAAC	TTCAAAAGTC	TGCCTAAGCT	1100
T Q S A	A D I R	K N F	K S L	P K L	
TGGGGCCGCA	CTGAGCACCC	ACCACCACCA	CCACTGAGAT	CCGGCTGCTA	1150
A A A	L E H H	H H H H	. D P A A N		
ACAAAGCCCG	AAAGGAAGCT	GAGTTGGCTG	GTGGCA		1186
K A R	K E A	E L A G	G		

FIG 9

10	20	30	40	50	
<hr/>					
1234567890	1234567890	1234567890	1234567890	1234567890	
<hr/>					
TGTCGGCTGT	GCTCCTGATC	CAGCACAGGC	GCCCATGGCC	TCTCCCAATT	50
C P L C S .	S S T G	A H C L	S Q L		
V R C A P D P	A Q A P I A	S P N W			
S A V L L I	Q H R R	P L P	L P I		
GGGCTAAAGG	CTTGGCATTG	TTCCTGCACA	GCTAAGTGCC	TGGGTTTCATC	100
G . R L A I V	P A Q L S A	W V H P			
A K G L P L	F L H S .	V P G F I			
G L K A C H C	S C T A K C L	G S S			
CTAATCGAGC	TGAACACTAG	TCACTGGGTT	CCACGGTTC	CTTCCATGAC	150
N R A E H .	S L G S T V L	F H D			
L I E L N T S	H W V P R F S	S M T			
. S S .	T L V T G F	H G S L P .	P		
CCATGGCTTC	TAATAGAGCT	ATAACACTCA	CTGCATGGTC	CAAGATTCCA	200
P W L L I E L .	H S L H G P	R F H			
H G F . .	S Y N T H	C M V Q D S I			
M A S N R A	I T L T A W S	K I P			
TTCCTTGGAA	TCCGTGAGAC	CAAGAACCCC	AGGTCAGAGA	ACACAAGGCT	250
S L E S V R P	R T P G Q R	T Q G L			
P W N P .	D Q E P Q	V R E H K A			
F L G I R E T	K N P R S E N	T R L			
TGOCACCATG	TTGGAAGCAG	CCCACCACCA	TTTTGGAAGC	AGCCCGCCAC	300
P P C W K Q	P T T I L E A	A R H			
C H H V G S S	P P P F W K Q	P A T			
A T M L E A A	H H H F G S	S P P L			
TATCTTGGGA	GCTCTGGGAG	CAAGGACCCC	AGGTACCAAT	TTGGTGAOCA	350
Y L G S S G S	K D P R .	Q F G D H			
I L G A L G A	R T P G N N	L V T T			
S W E L W E	Q G P Q V T I	W . P			
CGAAGGGACC	TGAATCCGCA	ACCATGAAGG	GATCTCCAAA	GCAATTGGAA	400
E G T .	I R N H E G	I S K A I G N			
K G P E S A	T M K G S P K	Q L E			
R R D L N P Q	P . R D L Q S	N W K			

FIG 9 (continuación)

10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	
ATGTTCTCTCC	CAAGGCAAAA	ATGCCCOCTAA	GATGTATTCT	GGAGAATTGG	450
V P P	K A K	M P L R	C I L	E N W	
M F L P	R Q K	C P .	D V F W	R I G	
C S S	Q G K N	A P K	M Y S	G E L G	
GACCAATTTG	ACCTCAGAC	AGTAAGAAAA	AAATGACTTA	TATTCTTCTG	500
D Q F D	P Q T	V R K K	. L I	F F C	
T N L	T L R Q	. E K	N D L	Y S S A	
P I .	P S D	S K K K	M T Y	I L L	
CAGTACCGOC	CTGGCCACGA	TATCCTCTTC	AAGGGGGAGA	AACTTGGCCT	550
S T A	L A T I	S S S	R G R	N L A S	
V P P	W P R	Y P L Q	G G E	T W P	
Q Y R P	G H D	I L F	K G E K	P G L	
OCTGAGGGAA	GTATAAATTA	TAACAACATC	TTACAGCTAG	ACCTGTTTTG	600
. G K	Y K L	. H H L	T A R	P V L	
P E G S	I N Y	N T I	L Q L D	L F C	
L R E V .	I I	T P S	Y S .	T C F V	
TAGAAAAGGA	GGCAAATGGA	GIGAAGTGCC	ATATTTACAA	ACTTTCTTTT	650
. K R R	Q M E	. S A	I F T N	F L F	
R K G	G K W S	E V P	Y L Q	T F F S	
E K E	A N G	V K C H	I Y K	L S F	
CATTAAAAGA	CAACTGGCAA	TTATGTTAAC	AGTGTGATT	GIGTTCTTAC	700
I K R	Q L A I	M L T V	. F V	F L H	
L K D	N S Q	L C .	Q C D L	C S Y	
H . K T	T R N	Y V N	S V I C	V P T	
AOGGAAGOC	TCAGATTCTA	CTCCCCAACC	CCGGCATCTC	CCCTGAATCC	750
G S P	Q I L	L P T P	G I S	P E S	
T E A L	R F Y	S P P	P A S P	L N P	
R K P	S D S T	P H P	R H L	P . I P	
CTCCCCAACC	TATT				764
L P N L					
S P T Y					
P Q L I					

FIG 10

10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	
TGTCGCTGTF	GCTCCIGATC	CAGCACAGGC	GCCCATGCCC	TCTCCCAATT	50
C P L C S .	S S T G A H C L	S Q L			
V R C A P D P	A Q A P I A	S P N W			
S A V L L I	Q H R R	P L P L P I			
GGGCTAAAGG	CTTGCCATTG	TTCCCTGCACA	GCTAAGTGGC	TGGGTTCATC	100
G . R L A I V	P A Q L S A	W V H P			
A K G L P L	F L H S .	V P G F I			
G L K A C H C	S C T A K C L	G S S			
CTAATCGAGC	TGAACACTAG	TCACTGGGTT	CCACGGTTCT	CTTCCATGAC	150
N R A E H .	S L G S T V L	F H D			
L I E L N T S	H W V P R F S	S M T			
. S S .	T L V T G F	H G S L P .	P		
CCATGGCTTC	TAATAGAGCT	ATAACACTCA	CTGCATGGTC	CAAGATTCCA	200
P W L L I E L .	H S L H G P	R F H			
H G F . .	S Y N T H C M V	Q D S I			
M A S N R A	I T L T A W S	K I P			
TTCCCTGGAA	TCGGTGAGAC	CAAGAACCCC	AGGTCAGAGA	ACACAAGGCT	250
S L E S V R P	R T P G Q R	T Q G L			
P W N P .	D Q E P Q	V R E H K A			
F L G I R E T	K N P R S E N	T R L			
TGCCACCATG	TTGGAAGCAG	CCCAACCACCA	TTTTGGAAGC	GGCCCCGCAC	300
P P C W K Q	P T T I L E A	A R H			
C H H V G S S	P P P F W K R	P A T			
A T M L E A A	H H H F G S	G P P L			
TATCTTGGGA	GCTCTGGGAG	CAAGGACCCC	CAGGTAACAA	TTTGGTGACC	350
Y L G S S G S	K D P Q V T I	W . P			
I L G A L G A	R T P R .	Q F G D H			
S W E L W E	Q G P P G N N	L V T			
ACGAAGGGAC	CTGAATCCGC	AACCATGAG	GGATCTCCAA	AGCAATGGGA	400
R R D L N P Q	P . R D L Q	S N W K			
E G T .	I R N H E G	I S K A I G			
T K G P E S A	T M K G S P K	Q L E			

ES 2 331 251 T3

FIG 10 (continuación)

10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	
AAATGTTCTC	CCAAGGCAAA	AAATGCCCTA	AGATGTATTG	TGGAGAATTG	450
C S S	Q G K	N A P K	M Y S	G E L	
N V P P	K A K	M P L	R C I L	E N W	
M F L	P R Q K	C P .	D V F	W R I G	
GGACCAATCT	GACCCICAGA	CAGTAAGAAA	AAAAATGACT	TATATTCTTC	500
G P I .	P S D	S K K	K N D L	Y S S	
D Q S	D P Q T	V R K	K M T	Y I L L	
T N L	T L R Q .	E K K .	L I F F		
TGCAGTACCG	OCTGGCCAGG	GATATCTCT	TCAAGGGGGA	GAAACCTGGC	550
A V P	P G H G	Y P L	Q G G	E T W P	
Q Y R	L A T	D I L F	K G E	K P G	
C S T A	W P R	I S S	S R G R	N L A	
CTCCTGAGGG	AAGTATAAAT	TATAACAACA	TCTTACAGCT	AGACCTGTTT	600
P E G	S I N	Y N T I	L Q L	D L F	
L L R E	V . I	I T P	S Y S .	T C F	
S . G	K Y K L .	H H	L T A	R P V L	
TGTAGAAAAG	GAGCCAAATG	GAGTGAAGTG	CCATATTTCAC	AAACTTICTT	650
C R K G	G K W	S E V	P Y L Q	T F F	
V E K	E A N G	V K C	H I Y	K L S F	
. K R	R Q M E .	S A	I F T	N F L	
TTCATTAAAA	GACAACCTGC	AATTATGTAA	ACAGTGTGAT	TGTGTGICCTA	700
S L K	D N S Q	L C K	Q C D	L C P T	
H . K	T T R	N Y V N	S V I	C V L	
F I K R	Q L A	I M .	T V .	F V S Y	
CAGGAAGGCC	TCAGATCTAC	CTCCCTAACC	GGCATCTCC	CTGACTCCTT	750
G S P	Q I Y	L P T P	A S P .	L L	
Q E A L	R S T	S L P	R H L P	D S F	
R K P	S D L P	P Y P	G I S	L T P S	
CCCCAATAA	TAGGACCCA	CTTCAGCCCA	AACAGTCCAA	AAGGACATAG	800
P Q L I	R T H	F S P	N S P K	G H	
P N .	. G P T	S A Q	T V Q	K D I	
P T N	K D P	L Q P K	Q S K	R T .	

FIG 11

10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	
GGCATTGATA	GCAOCCATCA	GATGGOCAAA	TCATTATTTA	CTGGACCAGG	50
G I D S	T H Q	M A K	S L F T	G P G	
A L I	A P I R	W P N	H Y L	L D Q A	
H . .	H P S	D G Q I	I I Y	W T R	
OCTTTTCAA	ACTATCAAGC	AGATAGGGCC	CGTGAAGCAT	GCCAAAGAAA	100
L F K	T I K Q	I G P	V K H	A K E I	
F S K	L S S R	. G P	. S M	P K K	
P F Q N	Y Q A	D R A	R E A C	Q R N	
TAATCCCCTG	CCTTATCGCC	ATGTTCCCTC	AGGAGAACAA	AGAACAGGCC	150
I P C	L I A	M F L Q	E N K	E Q A	
. S P A	L S P	C S F	R R T K	N R P	
N P L	P Y R H	V P S	G E Q	R T G H	
ATTACCCAGG	GGAAGACTGG	CAACTAGATT	TTACCCACAT	GGCCAAATGT	200
I T Q G	K T G N	. I	L P T W	P N V	
L P R	G R L A	T R F	Y P H	G Q M S	
Y P G	E D W	Q L D F	T H M	A K C	
CACGGATTTC	AGCATCTACT	AGTCTGGGCA	GATACTTTCA	CTGGTIGGGT	250
R D F	S I Y .	S G Q	I L S	L V G W	
G I S	A S T	S L G R	Y F H	W L G	
Q G F Q	H L L	V W A	D T F T	G W V	
GGAGICTTCT	CCTTGTAGGA	CAGAAAAGAC	CCAAGAGGTA	ATAAAGCCAC	300
S L L	L V G	Q K R P	K R .	. R H	
G V F S	L .	D R K D	P R G N	K G T	
E S S	P C R T	E K T	Q E V	I K A L	
TAATGAATA	ATTCCAGAT	TGGACTTCC	CCCAGGATTA	CAGGGTGACA	350
. . N N	S Q I	W T S	P R I T	G . Q	
N E I	I P R F	G L P	P G L	Q G D N	
M K .	F P D	L D F P	Q D Y	R V T	

FIG 11 (continuación)

10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	
ATGGCCCCGC	TTC AAGGCT	GCAGTAAACC	AGCGAGTATC	CCAGGIGTTA	400
W P R	F Q G C	S N P	G S I	P G V R	
G P A	F K A	A V T Q	G V S	Q V L	
M A P L	S R L	Q .	P R E Y P	R C .	
GCATACAAT	ATCACTTACA	CTGIGCCTGG	AGGCCACAAT	CCTOCAGAAA	450
H T I	S L T	L C L E	A T I	L Q K	
G I Q Y	H L H	C A W	R P Q S	S R K	
A Y N	I T Y T	V P G	G H N	P P E K	
AGICAAGAAA	ATGAATGAAA	CACTCAAAGA	TCTAAAAAAG	CTAACCCAAG	500
S Q E N	E .	N T Q R	S K K A	N P R	
V K K	M N E T	L K D	L K K	L T Q E	
S R K	. M K	H S K I	. K S	. P K	
AAACCCACAT	TGCATGACCT	GTICTGTTGC	CTATAACCTT	ACTAAGAATC	550
N P H	C M T C	S V A	Y N L	T K N P	
T H I A	. P V L L P	I T L	L R I		
K P T L	H D L	F C C L	. P Y	. E S	
CATACTATC	CCCCAAAAG	CAGGACTTAG	CCCATACGAG	ATGCTATAIG	600
. L S	P K K	Q D L A	H T R	C Y M	
H N Y P	P K S R T	. P I R D	A I W		
I T I	P Q K A	G L S	P Y E	M L Y G	
GATGGCCTTT	CCTAACCAAT	GACCTTGTGC	TTGACTGAGA	AATGGCCAAC	650
D G L S	. P M	T L C	L T E K	W P T	
M A F	P N Q .	P C A	. L R	N G Q L	
W P F	L T N	D L V L	D .	E M A N	
TTAGTIGCAG	ACATCACCTC	CTTAGCCAAA	TATCAACAAG	TTCTTAAAAC	700
. L Q	T S P P	. P N	I N K	F L K H	
S C R	H H L	L S Q I	S T S	S . N	
L V A D	I T S	L A K	Y Q Q	V L K T	

FIG 11 (continuación)

10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	
ATCACAGGGA	ACCTGTCCCC	GAGAGGAGGG	AAAGGAACTA	TGCCACCCIG	750
H R E	P V P	E R R E	R N Y	S T L	
I T G N	L S P	R G G	K G T I	P P W	
S Q G	T C P R	E E G	K E L	F H P G	
GTCACATG					758
V T					
. H					
D M					

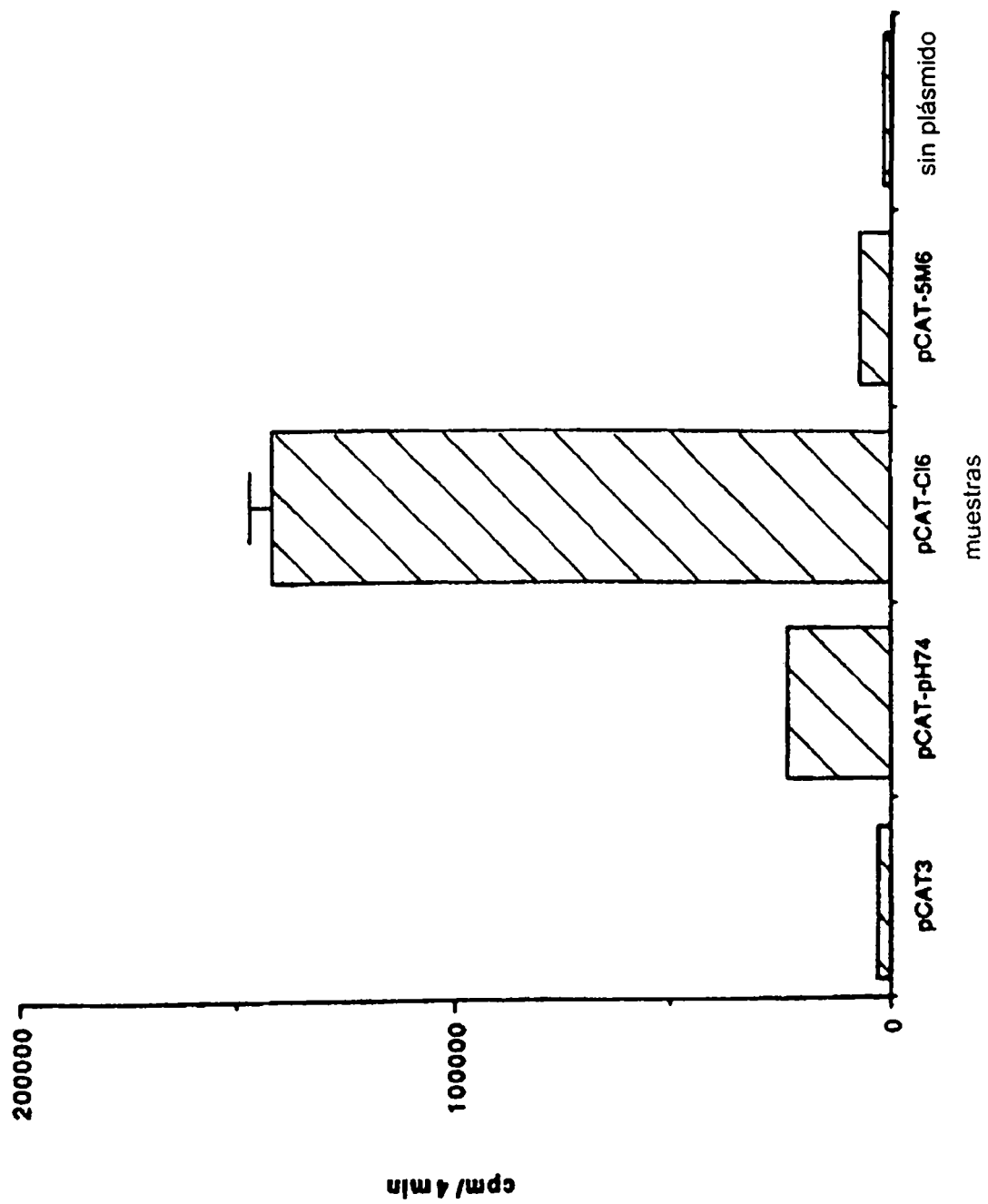


FIG 12

100 ATGGCCCTCC CTTATCTATAC TTCTCTCTTT ACTGTTCTCT TACCCTTTT TACCCTTACT GCGCTCTCTG CAGCCCTCTG TCAAGCCCTT AGCTCCCTCT
 34 K A L P Y M T P L P I Y L L L L P P P A L L I A P P P C C T T B S B P Y
 ACGAAGATT TCTATGAGCA ACCGCGCTTC CTGGAAATAT TCTATGCCCA TCAATAGACA GTTATATCTA GCGAACTCC ACTCTCACTG CCGACGCCA
 200 O E F L . R T R L P G M I D A P S Y R S L S K O P T A N T N
 67 TATGCCCGCC AACTCTATA ACTCTCTAGC TCTTGTGTC CARGAATA CTGATATG GAGCGGAA ATGATTTATC CTATCTCTTC TGGAGCACTT
 100 M P R M C Y M S A T L C M H A M T H Y W T G E M I N P S C P G G L
 GAGCCACTG TCTCTGTGAC TACTCTAGC CATACCAGTA TCTCTCTAGC GCGTCAAT CAAGCTGAGC AAGAGAAA ACAAGTAAAG GAGCCAACT
 400 G A T V C M T Y P T H T S N S D G G I Q G A R E E Q V E S A I S
 134 CCGAACTGAC CCGGGAAT AGCGCCGCA GCGCTTACAA AGCACTAATT CTCTCAAAAC TCACTGAAAC CTTCTGTAGC CATCTCTGCC TCTGTAGCCT
 300 Q L T R G H S T P S P Y K G L V L S K L H E T L R T N T R L V B L
 167 AITTAATCC AGCTCTACT GCGTCCATGA GCTCTGAGCC CAATAACCTA CTAACTGTG GATCTGCTC CCGCTCTACT TCAAGCCGTA CATTTCAATC
 600 P N T L T R L M E V S A Q N P T M C T M N C L P L M P R P Y I S I
 200 CCGTCTCTG AACAGCGGA GAACTCTAGC ACAGAAATA AACCACTTC CCGTTTATA GAACTCTTG TTTCGAATC GGAATGACC CATACTCTAA
 700 P V P E Q W M N P S T E I P T S V L V G P L V S M L E I T H T S P
 234 AGCTGACTG TGTAAATTT AGCAATACTA TAGACAGAC CAAGCTCGAA TCACTAGCCT GCGTACAGC TCCGACAGCA ATAGCTCTGC TACCTCTAGG
 800 C V K P S M T I D T T S E Q C I R M V T P P T R I V C L P S G
 367 AATATTTT GTCCTGCTA CCGTCAAGTA TCTCTCTC A A G E C C T T C A M A T A T G T C T C T C T T A T T G C C T A T T A C
 900 I P P V C C T S A Y H C L N G S S E S M C P L S P L V P P M T I Y
 300 ACTGACAG ATTATACAA TCACTGCTA CCGTAAAGCC ACAGCAAGAG AGTACCAATT CTCTCTTTTG TTATCAGAGC AGGAGCTA GCGACACTAG
 1000 T E O D L Y N M V V P K P H N K R V P I L P P V I R A O V L G R L G
 134 GTACTGAT TCGAGTATC ACAGCCGTA CTGAGTCTA CTAGAACTA TCTCAAGAA TCTCAAGAA TAAATGCTA CAGCGAAGC GTCACTACT CCGTCTCTAC
 1100 T G I G S I T T S T Q P Y Y K L S O E I N C D H E Q V T D S L V T
 367 CTCTCAAGAT CAACTTACT CCGTACAGC AGTACTCTTT CAATAAGAA GAGTTTACA CTCTCTTACC GCGAAAGAG GCGGAAGCTG TTATTTTAA
 1200 L O D Q L N S L A A V V L Q M R R A L D L L T A K R G G T C L P L
 400 GAGAGAAC CCGTGTATA TGTAAATCA TCCAGAAATC TCACTGACAA AGTTAAAGAA ATTGCAATC CAATCAATG TACAGCAGAG CAGCTTCAAA
 1100 G E R C Y Y V W U S R I V T E K V K E I R D K I Q C R A E S L Q M
 434 ACAGCAAGC CTGCGCCCTC CTGAGCCOAT GCACTGCTCC GCTTCTGCCC TTCTTAGCAC CTCTAGAGC TCTAATATTC TTACTCTCTT TTGAGCCCTG
 1400 T E R W G L L S Q W M P M V L P F L G P L A A L I L L L L P G P C
 467 TATLTTTAA CTTCTTGTATA AGTTTGTCTC TTCCAGAAIT GAACTGTAA ACCTACAGAT GCTTTTACAA ATGCAACCCC AGATGCAATC CATGACTTAAI
 1480 I P N L L V K P V S S R I E A V K L Q M V L Q M E P Q H E S M T K
 500 ATCAAGCTG GAGCCCTGCA CCGGCTCTCT AGCCGCTCT CCAATGTAA TCACTGTGAA GCGAGCCCTC CCGAGAAAT CTCAACTGCA CAAACCCCTAC
 1600 I H R G P L D R P A S P C S D V N D I E G T P P E E I S T A Q P L L
 534 TATGCCCGCA TCGAGCGGA AGCAATTA GCGTCTATCA GCGCACTCTC CCGAGGAC TTGCGTTTTC CTCTCTGACAG GCGGAGCTCA GAGCAGAGAC
 1700 C P N S A G S
 542 TAGCTGCAIT TCTTAGGCCA ACCAGAAATC CCGTAAAGCTA GCTGAGAGC TCACTGATC GAGCTCTAAA CARGGCGCTT GCAACTTACG TCAAGCCCA
 1800 CCAATCAGAG AGCTCACTAA AATCTCAAT AGCGAAAT AGCAATGAA CAATATGCA ATCACTTATT CCGTACAGC ACAGCCGAG GAGAGAGAT
 1900 CCGATATA ACCAGCCOAT TCGAGCGCC AACCGAAC CCGTTTGGCT TGTATGCCG CTCTGTTTTC ACTCTATTC ACTCTATTA
 2000 P T S V L V G P L V S M L E I T H T S P rely a signal
 2010 AACTTGGAC TGAAMAAA AAAAAAAAAA cap site

FIG13

FIG 14

CAGCAACC CTTGGGTC CTTCCCATG TAGGGAGCT CTGTTTCAC TCATTTCAC TCATTTCAC CATGCACTG CACTCTCTG GTCCGTTCT
 TTATGGCTC AGCTGAGCT TTTGTTGCC ATCCACACT GCTGTTGCC ACGTCCACG ACCGGCTCT GACTCCACT CATTGGATC CAGCAGAGG
 TCCGTTGTC TCTGANTCA GCACAGGCC CCATGGCTC TCCCAATGG GCTAAGGCT TGCATGTTT CTTGCACAGC TAGTGCCTG GGTTCATCT
 AACGAGTGC AACACTAGTC NTGGGTTCC ACGTTCTCT TCCATGACC ATGCGTTCTA ATAGAGCTAT AACACTACT GCATGTCCA AGATTCAT
 CCTTGGATC CGTGAGCCA AGAACCCGAG GTACAGAAC ACAGGCTTG CCACATGTT GGAAGCAGCC CACCACNTT TTGGAGCAG CCGCCACTA
 TCTTGGAGC TCTGGAGCA AGGACCCGAG GTACATGTT GGTGACCAAG AGGGACCTG AATCGGCAC CATGAGGGA TCTCCAAAGC ⁹⁸⁹ATCGCAAC
 3 H G N
 700 GTTCCCCCG AGGCAAAAT GCCCTAGAA CGTATCTGG AGATTTGGG CCAATGTGAC ACTCGAGCC TAGAAGAA ACGATTATA TTCTCTGCA
 37 V P P E A K M P L E R I L E N W D Q C D T Q T L R K K R K K R F I P F C S
 800 GTACCGCTG GCCCAATAT CCTCTTCAG GGAGAAAC CTGGTTCTT GAGGAATA TAAATTATA CATCATTTA CAGTAGACC TCTCTGTAG
 T A M P Q Y P L Q G R E T W L P E G S I N Y N I I L Q L D L F C R
 900 AAGGAGGC AANTGGAGT AATGCGTA TGTGCAACT TTCTTTTCAAT TAGAGACA CTCACAATA TGTAAAAGT GTGTTTATG CCTACAGGA
 103 K E G K W S E V P Y V Q T F P S L R D N S Q L C K K C G L C P T G
 1000 AGCCCTCAGA GTCCACTCC CTACCCAGC GTCCCTCC CCACTCTCTC CTCACTAT AAGGACCCC CTTTACCCA AAGGTCNA AAGGATAG
 S P Q S P P P Y P S V P S P T P S S T N K D P P L T Q T V Q K E I D
 1100 ACAAAGGCT AACCAATGA CCAAGAGT CCAATATCC CGGATATCC CCGTCCCAAG CAGTGAAGG AGGAGATTC GCGCCAGCA GAGTCCCTG
 K G V N N E P K S A N I P R L C P L Q A V R Q G E F G P A R V P V
 170 ACCTTTTCT CTCTCAGCT TAAAGCAAT TAAATAGAC CTAGTAAAT TCTCAGATA CCTCGAGCC TATATTGATG TTTTACAAAG GTTAGGACA
 1200 P F S L S D L K Q I K I D L G K F S D N P D G Y I D V L Q G L G Q
 203 TCCTTTGATC TGACATGGAG AGATATAATG TCACTACTA ATCAGACT AAGCCCAAT GAGAGAGT CCGCTGTAC TGCAGCCCA GAGTTTGGG
 S F D L T W R D I M L L L N Q T L T P N E R S A A V T A A R E F G D
 1400 ATCTTTGTA TCTCAGTCAG GCCAACATA GGATGACAC AGAGGAAGA ACACTCCA CAGGCCAGCA GCGAGTCCC AGTGTAGACC CTCAATGGGA
 L W Y L S Q A N N R M T T E E R T T P T G Q Q A V P S V D P H W D
 270 CACAGATCA GAACATGGAG ATGGTGCA CAACATTTG CTAAGTTGG TGTAGAGG ACTGAGGAA ACTAGGAGA AGCCTATGAA TTACTCAATG
 T E S E H G D W C H K H L L T C V L E G L R K K T R K K P M N Y S M
 303 ATGTCCACTA TACACAGGG AAGGAGAA AATCTTACTG CTTTCTGGA CAGACTAAG GAGCAATGA GGAAGCATC CTCCCTGTCA CCTCACTTA
 M S T I T Q G K E E N L T A F L D B L R E A L R K H T
 1600 TTGAAGCCA ACTAATCTTA AAGGATAGT TTATCACTA CTCAGTGA GACATTAGA AANAATCTA AAGTCCGTC TTAGGCTCG AACAAACTT
 1700 E G Q L I L K D K P I T Q S A A D I R K K L Q K S V L G S E Q N L
 370 AGAACCTTA TTGAAGTGG CAACTCGGT TTTTATAT AGAGTACAG AGGAGCAGC AGATGGGAG AANTGGATA AAAAAAAG GGCACCGCT
 E T L L N L A T S V P Y N R D Q E E Q A E W D K W D K K K R A T A
 403 TTAGTATGG CCTCAGGCA AGCGACTTT GGGGCTG GAAAGGCA AAGTGGCA ANTAGGAGC CTAAAGGC TTGCTTCAG TGGGTTCTAC
 1900 L V H A L R Q A D P G S G K G K S W / A N R K P N R A P Q G L Q
 437 AAGCACCT TAAANAAT TGTCCAAATA GAATAGCC GCCCTCTCT CCAATGCGCT TACGTCAAG GAATCACTGG AAGCCCACT GCGCCAGGG
 2000 G F K K D P N R N K P P P C / R P C P L R Q G N / H W K A H C P R G
 470 ATCAAGTAC TCTGAGTCAG AAGCCNTTA CCAAGTATC CAGGACAG ACTGA
 2055 S R Y S E S E A I N Q M I Q Q Q D
 487

FIG 15

100 GGACCCGTAG TATGGGGTAA TCCCTCCCGG GAUACCAAGC CCCCACTACTC AGAAGAGCA ATAGAAATGGG GAACCTCAGG AGGACATGOT TTCTCCCT
 34 G P V V W G H P L R E T K P Q Y S E E E I E M G T S R G H G P L P S
 200 CAGGATGGCT AGCCACTGA GAAGGAAJAA TACTTTTCT GCGAGCTAAC CAATOGAAAT TACTTAAJAC CTTTACGAA ACCTTCCACT TAGCCATICA
 67 G W L A T E E G K I L L L A A N Q W K L L K T L Q Q T P E L G I D
 300 TAGCAGCCAT CAGATAGCCA ATCATATT TACTGACCA GGCCTTTCA AJACTATCA GCAGATAGTC AGGGCTGTG AGTGTGCCA AJGAAATAT
 100 S T E Q I A K S L P T G P G L P K T I K Q I V R A E Z V E Q R M M
 400 CCGCTGCTT ATCGCAAGC TCTTTCAGGA GAAGAAJAA CAGCCAAITA CCGAAJAGAA GACTGGCAAC TAGATTTTAT CCAATGCCA AAATCAGCG
 134 P L P Y R Q A P S G E Q R T G N Y P R Z D W Q L D F I H M P K S Q G
 500 GATTTCACTG TCTACTACTC TGGTATGATA CTTTCACTGG TTGGGCAGAG GCTTCCCT GTAGGACAGA AAAGTTCCAA GAGGTATTA AGGCACTAGT
 167 F Q C L L V W V D T P T G W A E A P P C R T E K P Q E V I X A L V
 600 TCATGAAGTA ATTCACAGT TCGACTCC CTGAGCTTA CAGAGTGACA ATGCTCTGC TTTCJAGGCC ACAGTAAACC AGGCACTATC CCAGGCTTA
 200 H E V I P R P Q L P . G L Q S D N G P A P K A T V T Q G V S Q A L
 700 GGTATAGAT ATCACTTACA CTGCACCTAG AGGCCAAT CTTGAGGAA GCTTACAGAA ATCAJACAC TCAJAGACA TCTTAAJAG CTAAJCCAGG
 234 G I E Y H L H C T . R P Q S S G K V E K M K T L K R H L M K L T Q E
 800 AAJCCACTT CCGATGCTT GCTCTGTTCT CTATAGCTT ACTAAGATC CAJACTCTC CCAJAGGC AGCACTTAGC CCAJAGCAA TCTGTATAG
 267 T H L A M S A L L S I A L L R I Q M S P Q K A G L S P Y R M L Y G
 900 AGGTCTTC CTAAJCAJG ACCTTCTGCT TCAJCAJAG ATGGCCACT TAGTTGACA CATCACTCC TTAGCCAAAT ATCAJCAJG TCTTAAJACA
 300 R S P L T M D L L L D Q E M A M L V A D I T S L A K Y Q Q V L E T
 1000 TTACAJAG CTTGTCCCG AGAGAGGGA AJAGAAATF TCCJCCCTG TGTCAJGTA TTAGTCAJG TCCCTTCTC TAAJTCOCA TCCCTAGACA
 334 L Q G A C P R Z E G K E I P H P G V M V L V K S L P S M S P S L D T
 1100 CAJCTGGGG AGJACCTTAC CCAJCTATT TAJCTATCC AJCTCGCTT AJJGTGCTG GAGTGCAGTC TTGATACAT CAJACTGAA TCAJACCTG
 367 S W G G P Y P V I L S I P T A V K V A G V E S W I M H T R I K P W
 1197 GATACTCG AJGAAJCCG AJAATCCAGG GCAJAGCT AJCTATTCT TTCAJCTT AGAJCATCTC AGAJCATCTC TCCCTGCTT TCAJAGCA AJCCTGA
 396 I L P K E P E N P G D N A S Y P F E P L E D L C L L P K Q Q P

FIG 16

100 OAGREBGA OGTUAGTGS OGTREBGA OGTREBANG AKOAGAGA OGTANAGAA MAAGAGAA OGTAGANA GANANAGA OGTAGANA
 E N S S I S W L A E V G K D S K K R K K O E S Q R K K E R E Z E T
 200 OAGAGAGA OGTAGAGA OAGAGABA OGTAGAGA MAAGAGAG OGTREBCTT BANAGANG OGTANAGTIC TOGTAGTIA OGTAGAGNE
 K K N L K R E R S S K E K T V Y P I P L K A R V N F C L P S Q G I
 300 XTTTCTTA TGTAGAT OAGCTTTC OGTREBGA OGTREBANG MAAGTGTIC TTTTGTAGT OGTAGTGA OGTAGTGA OGTAGTGA
 F P L C G T S T Y I C L P T N W T G T R T L V F L S P N I N I A P
 400 OAGAGAGA OGTAGTGT OGTAGTGA OGTAGTGT OGTAGTGA MAAGAGAGA OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT
 G N O T L L V P V K A K V R O C H A I Q L I S L F I G L G M A T A T
 500 OAGAGTGS MAAGTGT TGTAGTGT OGTAGTGA OGTAGTGA OGTAGTGA OGTAGTGA OGTAGTGA OGTAGTGA OGTAGTGT
 G T G I A G L S T S L S Y Y K T L S K N F S D S L O E I M K S I L
 600 TACTTTCGA TGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT
 T L Q S Q L D S L A A H T L Q N R R G P H L L L T A E K G L C T P
 700 TTAGAGAG MAAGTGT TGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT
 L G P C P Y T N O S G I V R D A T W H L Q E R A S D I R Q C L S
 800 OAGTGTGA TGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT
 N S Y T H L W S W A T W L L P F L G P N A A I L L L L T P G P C I
 900 TTTAGTGT OGTAGTGT TGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT
 F K L L V K F V S S R I E A I K L Q H V L Q H E P Q H S S T N N F
 1000 TTAGAGAG OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT
 Y Q G P L E R S T C T S T S L E I P L W K T L Q L Q G P F F A P I Q
 1100 OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT
 Q E V A R A V I G O I P N S S W G V L F R G G I E E A C W Q P
 1200 TTAGTGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT
 K S P R M I S V P P Q P W C P L W P C L R S P S A C H C T V G A S
 1300 TTTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT
 F W A G Q G R S Q L P Q L A G R Y G G R D A G G N Q G C A W R L R A
 1400 OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT
 S H B S R N A W A R R A P H S G S E C L S T W A R Q H L C S T S S
 1500 OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT
 L G L S C L P R G A G L R E H A A C P C L S P P P R R G F L H S P
 1600 OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT
 S P P D K H K P L S T V P S P I N H P R V E E C H T A R D W Q A V
 1700 TTTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT OGTAGTGT
 P L A A L V R D P L R E A S W A P E S G G D L E N L Y V J H R D C
 TTAGTGTGT OGTAGTGT
 K Y T N Q H

ES 2 331 251 T3

LISTA DE SECUENCIAS

	<110> bioMérieux	
5	<120> Material nucleico retroviroico y fragmentos nucleotidicos, asociados a la esclerosis en placas, con fines de diagnóstico, profilácticos y terapéuticos	
	<130> BR51193	
	<140>	
10	<141>	
	<150> FR-9708816	
	<151> 07-07-1997	
15	<160> 31	
	<170> PatentIn Ver. 2.1	
	<210> 68	
20	<211> 34	
	<212> ADN	
	<213> humano	
25	<400> 68	
	gactcgctgc agatcgattt tttttttt tttt	34
30	<210> 69	
	<211> 30	
	<212> ADN	
35	<213> humano	
	<400> 69	
40	gccatcaagc cacccaagaa ctcttaact	30
	<210> 70	
	<211> 30	
45	<212> ADN	
	<213> humano	
	<400> 70	
50	ccaatagcca qadeattata tacactaatt	30
	<210> 112	
55	<211> 310	
	<212> ADN	
	<213> humano	
60		
65		

ES 2 331 251 T3

<400> 112

5 gcttatagaa ggacccttag tatggggtaa tcccctctgg gaaaccaagc cccagtactc 60
 agcaggaaaa atagaatagg aaacctcaca aggacataact ttctcccct ccagatggct 120
 agccactgag gaaggaaaaa tactttcacc tgcagctaac caacagaaat tacttaaaac 180
 ccttcaccaa accttccact taggcattga tagcacccat cagatggcca aattattatt 240
 10 tactggacca ggccttttca aaactatcaa gaagatagtc aggggctgtg aagtgtgcca 300
 aagaaataat 310

<210> 113

15 <211> 103

<212> PRT

<213> humano

20 <400> 113

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

<210> 114

<211> 635

<212> ADN

50 <213> humano

<400> 114

55 ccctgtatct ttaacctctt tgtaagttt gtctcttcca gaatcaaaac tqtaaaacta 60
 caaattgttc ttcaaatgga gcaccagatg gagtccatga ctaagatcca ccgtggacc 120
 ctggaccggc ctgctagccc atgctccgat gttaatgaca ttgaaggcac ccctcccag 180
 gaaatctcaa ctgcacaacc cctactatgc cccaattcag cgggaagcag ttagagcgg 240
 60 catcagccaa cctcccac acgacttggg ttttctgtt gagagggggg actgagagac 300
 aggactagct ggatttccta ggccaacgaa gaatccctaa gcctagctgg gaaggtgact 360
 gcatccacct ctaaaccatg ggcttgcaac ttagctcaca cccgaccaat cagagagctc 420
 actaaaatgc taattaggca aaaataggag gtaaagaat agccaatcat ctattgcctg 480
 agagcacagc gggagggaca aggatcggga tataaaccca ggcattcgag cgggcaacgg 540
 65 caacccccct tgggtcccct ccctttgtat gggcgctctg ttttactct atttactct 600
 attaaatctt gcaactgaaa aaaaaaaaaa aaaa 635

ES 2 331 251 T3

<210> 115

<211> 77

<212> PRT

5 <213> humano

<400> 115

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

<210> 116

<211> 32

30 <212> ADN

<213> humano

<400> 116

35

tggggtcca ttgtaagac catctgtac tt

32

<210> 117

40 <211> 1481

<212> ADN

<213> humano

45

50

55

60

65

ES 2 331 251 T3

<400> 117

5 atggccctcc cttatcatac tttctctctt actgtctctt taccaccttt cgctctcact 60
 gcacccccctc catgctgctg tacaaccagt agctccccctt accaagagtt tctatgaaga 120
 acgcggcttc ctggaataat tgatgcccc a catatagga gtttatctaa gggaaactcc 180
 accttcaactg cccacaccca tatgccccgc aactgctata actctgccac tctttgcatg 240
 catgcaaata ctcattattg gacagggaaa atgattaatc ctagtgtgcc tggaggactt 300
 10 ggagccactg tctgttggac ttractcacc cataccagta tgtctgatgg ggttggatt 360
 caaggtcagg caagagaaaa acaagttaa ggaagcaatct cccaactgac ccggggacat 420
 agcaccctta gccctacaa aggactagtt ctctcaaaac tacatgaaac cctccgtacc 480
 catactcgcc tgggtgagcct atttaatacc accctcactc ggctccatga ggtctcagcc 540
 15 caaaacccta ctaactggtg gatgtgcctc cccctgcact tcaggccata cttttcaatc 600
 cctgttctctg aacaatggaa caacttcagc acagaaataa acaccacttc cgttttagta 660
 ggacctcttg tttccaatct ggaaataacc catacctcaa acctcacctg tgtaaaattt 720
 agcaatacta tagacacaac cagctcccaa tgcactcagg gggtaacacc tcccacacga 780
 20 atagtctgcc taccctcagg aatatttttt gtctgtggta cctcagccta tcattgtttg 840
 aatggctctt cagaatctat gtgcttctct cctattcttag tgcacctat gacctctac 900
 actgaacaag atttatacaa tcatgtctga cctaagcccc acaacaaaag agtaccatt 960
 ctctcttttg ttatcagagc aggagtctga ggcagactag gtactggcat tggcagtatc 1020
 25 acaacctcta ctcagttcta ctacaaacta tctcaagaaa taaatggtga catggaacag 1080
 gtcactgact ccttggctac cttgcaagat caacttaact ccctagcagc agtagtctct 1140
 caaatcgaa gagctttaga cttgctaacc gccaaaagag ggggaacctg tttattttta 1200
 ggagaagaac gctgttatta tgttaataca tccagaattg tcaactgagaa agttaaagaa 1260
 30 attcagatc gaatacaatg tagagcagag gagcttcaaa acaccgaacg ctggggcctc 1320
 ctcagccaat gtagccctg ggttctcccc ttcttaggac ctctagcagc tctaatactg 1380
 ttractctct ttggacctg tatctttaac ctcttgtta agtttctctc tccagaatt 1440
 gaagctgtaa agctacagat ggtcttcaaa atggaacccc a 1481

35 <210> 118
 <211> 493

40 <212> PRT
 <213> humano

<400> 118

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

ES 2 331 251 T3

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

ES 2 331 251 T3

	325	330	335
5	Ile Gly Ser Ile Thr Thr Ser Thr Gln Phe Tyr Tyr Lys Leu Ser Gln		
	340	345	350
10	Glu Ile Asn Gly Asp Met Glu Gln Val Thr Asp Ser Leu Val Thr Leu		
	355	360	365
15	Gln Asp Gln Leu Asn Ser Leu Ala Ala Val Val Leu Gln Asn Arg Arg		
	370	375	380
20	Ala Leu Asp Leu Leu Thr Ala Lys Arg Gly Gly Thr Cys Leu Phe Leu		
	385	390	400
25	Gly Glu Glu Arg Cys Tyr Tyr Val Asn Gln Ser Arg Ile Val Thr Glu		
	405	410	415
30	Lys Val Lys Glu Ile Arg Asp Arg Ile Gln Cys Arg Ala Glu Glu Leu		
	420	425	430
35	Gln Asn Thr Glu Arg Trp Gly Leu Leu Ser Gln Trp Met Pro Trp Val		
	435	440	445
40	Leu Pro Phe Leu Gly Pro Leu Ala Ala Leu Ile Leu Leu Leu Leu Phe		
	450	455	460
45	Gly Pro Cys Ile Phe Asn Leu Leu Val Lys Phe Val Ser Ser Arg Ile		
	465	470	475
50	Glu Ala Val Lys Leu Gln Met Val Leu Gln Met Glu Pro		
	485	490	

55 <210> 119
 <211>32
 <212> ADN
 <213> humano
 60 <400> 119

tcaaaatcga agagctttag acttgctaac cg

32

65 <210> 120
 <211> 1329

ES 2 331 251 T3

<212> ADN

<213> humano

5 <400> 120

tcaaaaatcga agagctcttag acttgctaac cgccaaaaga gggggaacct gtttatcttt 60
 aggggaagaa tgctgttagc atgttaatca atctggaatc attactgaga aagttaaaga 120
 10 aattcgagat cgaatataat gtagagcaga ggaccttcaa aacactgcac cctggggcct 180
 cctcagccaa tggatgccct ggactctccc ctctcttagga cctctagcag ctataatatt 240
 tttactcctc tttggaccct gtatcttcaa ctctcttgtt aagtttgtct cttccagaat 300
 15 tgaagctgta aagctacaaa tagttcttca aatggaacct cagatgcagt ccatgactaa 360
 aatctaccgt ggacccttgg accggcctgc tagactatgc tctgatgta atgacattga 420
 agtcaccctt cccgaggaaa tctcaactgc acaaccctta ctacactcca attcagtagg 480
 aagcagttag agcagttgtc agccaacctc cccaacagta ctgggtttt cctgttgaga 540
 20 ggggtggactg agagacagga cttagctggat ttcttaggct gactaagaat ccnaagcct 600
 anctgggaag gtgaccgcat ccatctttaa acatggggct tgcaacttag ctcacaccctg 660
 accaatcaga gagctcacta aaatgctaata caggcaaaaa caggaggtaa agcaatagcc 720
 25 aatcatctat tgccctgagag cacagcggga aggacaagga ttgggatata aactcaggca 780
 ttcaagccag caacagcaac cccctttggg tccccctcca ttgtatggga gctctgtttt 840
 cactctatct cactctatta aatcatgcaa ctgcaactct ctggtccgtg ttttttatgg 900
 ctcaagctga gcttttgttc gccatccacc actgctgttt gccaccgtca cagaccctgct 960
 30 gctgacttcc atccctttgg atccagcaga gtgtccactg tgctcctgat ccagcgaggt 1020
 acccattgcc actcccgatc aggctaaagg cttgccattg ttcttgcatt gctaagtgcc 1080
 tgggtttgtc ctaatagaac tgaacactgg tcaactgggtt ccatgggtct cttccatgac 1140
 35 ccacggcttc taatagagct ataacactca ccgcatggcc caagattcca ttctttggta 1200
 tctgtgaggc caagaacccc aggtcagaga angtgaggct tgccaccatt tgggaagtgg 1260
 cccactgcca ttttggtagc ggcccaccac catcttggga gctgtgggag caaggatccc 1320
 ccagtaaca 1329

40

<210> 121

<211> 162

<212> PRT

45

<213> humano

50

55

60

65

ES 2 331 251 T3

<400> 121

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

<210> 122

<211> 21

45 <212> ADN

<213> humano

<400> 122

50 ggcattgata gcacccatca g

21

<210> 123

55 <211> 21

<212> ADN

<213> humano

60 <400> 123

catgtcacca ggggtgaata g

21

65 <210> 124

<211> 758

<212> ADN

ES 2 331 251 T3

<213> humano

<400> 124

5 **ggcattgata gcacccatca gatggccaaa tcattathta ctggaccagg ccttttcaaa 60**
actatcaagc agatagggcc cgtgaagcat gccaaagaaa taatcccctg ccttatcgcc 120
atgttccttc aggagaacaa agaacaggcc attaccaggg ggaagactgg caactagatt 180
ttacccacat ggccaaatgt cagggathte agcatctact agtctgggca gatactttca 240
10 **ctgggtgggt ggagctcttc ccttqtagga cagaaaagac ccaagaggt aataaggcac 300**
taatgaaata attcccagat ttggacttcc cccaggatta cagggtgaca atggccccgc 360
tttcaagget gcagtaacce agggagtatc ccagggtgta ggcatacaat atcacttaca 420
15 **ctgtgcctgg aggccacaat cctccagaaa agtcaagaaa atgaatgaaa cactcaaaga 480**
tctaaaaaag ctaacccaag aaacccacat tgcattgacct gttctgttgc ctataacctt 540
actaagaatc cataactatc ccccaaaaag caggacttag cccatacgag atgctatatg 600
20 **gatggccttt cctaaccaat gaccttgtgc ttgactgaga aatggccaac ttagttagcag 660**
acatcacctc cttagccaaa tateacaag ttcttaaac atcacagggg acctgtcccc 720
gagaggaggg aaaggaacta tccaccctg gtgacatg 758

25 <210> 126

<211> 25

<212> ADN

<213> humano

30

<400> 126

eggacatcca aagtgatggg aaacg

25

35

<210> 127

<211> 26

<212> ADN

40

<213> humano

<400> 127

45

ggacaggaaa gtaagactga gaagc

26

<210> 128

<211> 26

50

<212> ADN

<213> humano

<400> 128

55

cctagacgt attctggaga attggg

26

<210> 129

60

<211> 26

<212> ADN

<213> humano

65

<400> 129

tgctctcaa tggccaaca taccgg

26

ES 2 331 251 T3

<210> 130
 <211> 1511
 <212> ADN
 5 <213> humano
 <400> 130

```

10      cctagaacgt attctggaga attgggacca atgtgacact cagacgctaa gaaagaacg 60
      atttatattc ttctgcagta ccgcctggcc acaatatacct cttcaaggga gagaaacctg 120
      gcttcctgag ggaagtataa attataaca? catcttacag cttagacctcc tctgtagaaa 180
      ggaggggcaaa tggagtgaag tgccatattg gcaaaccttc ttttcattaa gagacaactc 240
15      acaattatgt aaaaagtgtg gtttatgcc tacaggaagc cctcagagtc cacctcccta 300
      ccccagcgtc cctcctccga ctcttctctc aactaataag gacccccctt taacccaac 360
      ggtccaaaag gagatagaca aaggggtaaa caatgaacca aagagtgcc aatattccccg 420
      attatgcccc ctccaagcag tgagaggagg agaattcggc ccagccagag tgcctgtacc 480
20      tttttctctc tcagacttaa agcaaatata aatagacctt ggtaaattct cagataaccc 540
      tgacggctat attgatgttt tacaaggggt aggacaatcc tttgatctga catggagaga 600
      tataatgtta ctactaaac agacaactaac cccaaatgag agaagtgccg ctgtaactgc 660
      agcccagag tttggcgatc tttggtatct cagtcaggcc aacaatagga tgacaacaga 720
      ggaaagaaca actcccacag gccagcaggc agttcccagt gtagacctc attgggacac 780
25      agaatcagaa catggagatt ggtgccacaa acatttgcta acctgcgtgc tagaaggact 840
      gaggaaaact aggaagaagc ctatgaatta ctcaatgatg tccactataa cacagggaaa 900
      ggaagaaaat cttactgctt tcttgacag actaaggag gcattgagga agcatacctc 960
      cctgtcacct gactctattg aaggccaact aatcttaag gataaqtta tcactcagtc 1020
30      agctgcagac attagaaaaa acttcaaaag tctgccttag gcccgagca gaacttagaa 1080
      accctatta acttgcatc ctctgctttt tataatagag atcaggagga gcaggcga 1140
      cgggacaaac ggataaaaa aaaaaggggg ggtccactac tttagtcatg gccctcaggc 1200
      aagcagactt tggaggctct gcaaaagggg aaagctgggc aaatcaaatg cctaataagg 1260
      ctggcttcca gtgcggctca caaggacact ttaaaaaaga ttatccaagt agaataagc 1320
35      cgcctcttg tccatgcccc ttactgcaag ggaatcactg gaagggccac tgccccaggg 1380
      gatgaagata ctctgagtca gaagccatta accagatgat ccagcagcag gactgagggt 1440
      gcccgggcg agcggcagcc catgccatca cctcacaga gccccgggta tgtttgacca 1500
      ttgagagcca a                                     1511
  
```

<210> 131
 <211> 352
 <212> PRT
 45 <213> humano
 <400> 131

```

50      Leu Glu Arg Ile Leu Glu Asn Trp Asp Gln Cys Asp Thr Gln Thr Leu
           1             5             10            15

55      Arg Lys Lys Arg Phe Ile Phe Phe Cys Ser Thr Ala Trp Pro Gln Tyr
           20             25            30

60      Pro Leu Gln Gly Arg Glu Thr Trp Leu Pro Glu Gly Ser Ile Asn Tyr
           35             40            45

65      Asn Ile Ile Leu Gln Leu Asp Leu Phe Cys Arg Lys Glu Gly Lys Trp
           50             55            60
  
```

ES 2 331 251 T3

Ser Glu Val Pro Tyr Val Gln Thr Phe Phe Ser Leu Arg Asp Asn Ser
 65 70 75 80
 5
 Gln Leu Cys Lys Lys Cys Gly Leu Cys Pro Thr Gly Ser Pro Gln Ser
 85 90 95
 10
 Pro Pro Pro Tyr Pro Ser Val Pro Ser Pro Thr Pro Ser Ser Thr Asn
 100 105 110
 15
 Lys Asp Pro Pro Leu Thr Gln Thr Val Gln Lys Glu Ile Asp Lys Gly
 115 120 125
 20
 Val Asn Asn Glu Pro Lys Ser Ala Asn Ile Pro Arg Leu Cys Pro Leu
 130 135 140
 25
 Gln Ala Val Arg Gly Gly Glu Phe Gly Pro Ala Arg Val Pro Val Pro
 145 150 155 160
 30
 Phe Ser Leu Ser Asp Leu Lys Gln Ile Lys Ile Asp Leu Gly Lys Phe
 165 170 175
 35
 Ser Asp Asn Pro Asp Gly Tyr Ile Asp Val Leu Gln Gly Leu Gly Gln
 180 185 190
 40
 Ser Phe Asp Leu Thr Trp Arg Asp Ile Met Leu Leu Leu Asn Gln Thr
 195 200 205
 45
 Leu Thr Pro Asn Glu Arg Ser Ala Ala Val Thr Ala Ala Arg Glu Phe
 210 215 220
 50
 Gly Asp Leu Trp Tyr Leu Ser Gln Ala Asn Asn Arg Met Thr Thr Glu
 225 230 235 240
 55
 Glu Arg Thr Thr Pro Thr Gly Gln Gln Ala Val Pro Ser Val Asp Pro
 245 250 255
 60
 His Trp Asp Thr Glu Ser Glu His Gly Asp Trp Cys His Lys His Leu
 260 265 270
 65
 Leu Thr Cys Val Leu Glu Gly Leu Arg Lys Thr Arg Lys Lys Pro Met
 275 280 285
 70
 Asn Tyr Ser Met Met Ser Thr Ile Thr Gln Gly Lys Glu Glu Asn Leu
 290 295 300
 75
 Thr Ala Phe Leu Asp Arg Leu Arg Glu Ala Leu Arg Lys His Thr Ser
 305 310 315 320

ES 2 331 251 T3

Leu Ser Pro Asp Ser Ile Glu Gly Gln Leu Ile Leu Lys Asp Lys Phe
 325 330 335

5 Ile Thr Gln Ser Ala Ala Asp Ile Arg Lys Asn Phe Lys Ser Leu Pro
 340 345 350

10 <210> 132
 <211> 30
 <212> ADN
 <213> humano
 15 <400> 132

20 tgctggaatt cgggatccta gaacgtattc 30

<210> 133
 <211> 30
 <212> ADN
 <213> humano
 25 <400> 133

30 agttctgctc cgaagcttag gcagactttt 30

<210> 135
 <211> 398
 <212> PRT
 <213> humano
 35 <400> 135

Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val Pro
 1 5 10 15

Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg
 20 25 30

Ile Leu Glu Arg Ile Leu Glu Asn Trp Asp Gln Cys Asp Thr Gln Thr
 35 40 45

Leu Arg Lys Lys Arg Phe Ile Phe Phe Cys Ser Thr Ala Trp Pro Gln
 50 55 60

Tyr Pro Leu Gln Gly Arg Glu Thr Trp Leu Pro Glu Gly Ser Ile Asn
 65 70 75 80

ES 2 331 251 T3

Tyr Asn Ile Ile Leu Gln Leu Asp Leu Phe Cys Arg Lys Glu Gly Lys
 85 90 95
 5
 Trp Ser Glu Val Pro Tyr Val Gln Thr Phe Phe Ser Leu Arg Asp Asn
 100 105 110
 10
 Ser Gln Leu Cys Lys Lys Cys Gly Leu Cys Pro Thr Gly Ser Pro Gln
 115 120 125
 15
 Ser Pro Pro Pro Tyr Pro Ser Val Pro Ser Pro Thr Pro Ser Ser Thr
 130 135 140
 20
 Asn Lys Asp Pro Pro Leu Thr Gln Thr Val Gln Lys Glu Ile Asp Lys
 145 150 155 160
 25
 Gly Val Asn Asn Glu Pro Lys Ser Ala Asn Ile Pro Arg Leu Cys Pro
 165 170 175
 30
 Leu Gln Ala Val Arg Gly Gly Glu Phe Gly Pro Ala Arg Val Pro Val
 180 185 190
 35
 Pro Phe Ser Leu Ser Asp Leu Lys Gln Ile Lys Ile Asp Leu Gly Lys
 195 200 205
 40
 Phe Ser Asp Asn Pro Asp Gly Tyr Ile Asp Val Leu Gln Gly Leu Gly
 210 215 220
 45
 Gln Ser Phe Asp Leu Thr Trp Arg Asp Ile Met Leu Leu Leu Asn Gln
 225 230 235 240
 50
 Thr Leu Thr Pro Asn Glu Arg Ser Ala Ala Val Thr Ala Ala Arg Glu
 245 250 255
 55
 Phe Gly Asp Leu Trp Tyr Leu Ser Gln Ala Asn Asn Arg Met Thr Thr
 260 265 270
 60
 Glu Glu Arg Thr Thr Pro Thr Gly Gln Gln Ala Val Pro Ser Val Asp
 275 280 285
 65
 Pro His Trp Asp Thr Glu Ser Glu His Gly Asp Trp Cys His Lys His
 290 295 300
 Leu Leu Thr Cys Val Leu Glu Gly Leu Arg Lys Thr Arg Lys Lys Pro
 305 310 315 320
 Met Asn Tyr Ser Met Met Ser Thr Ile Thr Gln Gly Lys Glu Glu Asn
 325 330 335

ES 2 331 251 T3

Leu Thr Ala Phe Leu Asp Arg Leu Arg Glu Ala Leu Arg Lys His Thr
 340 345 350

5 Ser Leu Ser Pro Asp Ser Ile Glu Gly Gln Leu Ile Leu Lys Asp Lys
 355 360 365

10 Phe Ile Thr Gln Ser Ala Ala Asp Ile Arg Lys Asn Phe Lys Ser Leu
 370 375 380

15 Pro Lys Leu Ala Ala Ala Leu Glu His His His His His His
 385 390 395

<210> 137

20 <211> 378

<212> PRT

<213> humano

25 <400> 137

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

30 Ile Leu Glu Asn Trp Asp Gln Cys Asp Thr Gln Thr Leu Arg Lys Lys
 20 25 30

35 Arg Phe Ile Phe Phe Cys Ser Thr Ala Trp Pro Gln Tyr Pro Leu Gln
 35 40 45

40 Gly Arg Glu Thr Trp Leu Pro Glu Gly Ser Ile Asn Tyr Asn Ile Ile
 50 55 60

45 Leu Gln Leu Asp Leu Phe Cys Arg Lys Glu Gly Lys Trp Ser Glu Val
 65 70 75 80

50 Pro Tyr Val Gln Thr Phe Phe Ser Leu Arg Asp Asn Ser Gln Leu Cys
 85 90 95

55 Lys Lys Cys Gly Leu Cys Pro Thr Gly Ser Pro Gln Ser Pro Pro Pro
 100 105 110

Tyr Pro Ser Val Pro Ser Pro Thr Pro Ser Ser Thr Asn Lys Asp Pro
 115 120 125

60 Pro Leu Thr Gln Thr Val Gln Lys Glu Ile Asp Lys Gly Val Asn Asn
 130 135 140

65 Glu Pro Lys Ser Ala Asn Ile Pro Arg Leu Cys Pro Leu Gln Ala Val

ES 2 331 251 T3

	145		150		155		160
5	Arg Gly Gly Glu Phe Gly Pro Ala Arg Val Pro Val Pro Phe Ser Leu						
		165		170		175	
10	Ser Asp Leu Lys Gln Ile Lys Ile Asp Leu Gly Lys Phe Ser Asp Asn						
		180		185		190	
15	Pro Asp Gly Tyr Ile Asp Val Leu Gln Gly Leu Gly Gln Ser Phe Asp						
		195		200		205	
20	Leu Thr Trp Arg Asp Ile Met Leu Leu Leu Asn Gln Thr Leu Thr Pro						
		210		215		220	
25	Asn Glu Arg Ser Ala Ala Val Thr Ala Ala Arg Glu Phe Gly Asp Leu						
		225		230		235	
30	Trp Tyr Leu Ser Gln Ala Asn Asn Arg Met Thr Thr Glu Glu Arg Thr						
		245		250		255	
35	Thr Pro Thr Gly Gln Gln Ala Val Pro Ser Val Asp Pro His Trp Asp						
		260		265		270	
40	Thr Glu Ser Glu His Gly Asp Trp Cys His Lys His Leu Leu Thr Cys						
		275		280		285	
45	Val Leu Glu Gly Leu Arg Lys Thr Arg Lys Lys Pro Met Asn Tyr Ser						
		290		295		300	
50	Met Met Ser Thr Ile Thr Gln Gly Lys Glu Glu Asn Leu Thr Ala Phe						
		305		310		315	
55	Leu Asp Arg Leu Arg Glu Ala Leu Arg Lys His Thr Ser Leu Ser Pro						
		325		330		335	
60	Asp Ser Ile Glu Gly Gln Leu Ile Leu Lys Asp Lys Phe Ile Thr Gln						
		340		345		350	
65	Ser Ala Ala Asp Ile Arg Lys Asn Phe Lys Ser Leu Pro Lys Leu Ala						
		355		360		365	
70	Ala Ala Leu Glu His His His His His His						
		370		375			

ES 2 331 251 T3

<210> 138
 <211> 25
 <212> ADN
 5 <213> humano

 <400> 138

 10 cttqgagggt geataaccag ggaat 25

 <210> 139
 <211> 20
 15 <212> ADN
 <213> humano

 <400> 139
 20 tgtccgctgt gctcctgatc 20

 <210> 140
 25 <211> 25
 <212> ADN
 <213> humano

 30 <400> 140

 ctatgctcctt ttggactggt tgggt 25

 35 <210> 141
 <211> 764
 <212> ADN
 40 <213> humano

 <400> 141

 tgcccgctgt gctcctgatc cagcacaggc gccattgce tctcccaatt gggctaaagg 60
 45 **cttgccattg ttcctgcaca gctaagtgcc tgggttcate ctaatcgagc tgaacactag 120**
 tcactgggtt ccacggttct ctccatgac ccatggcttc taatagagct araacactca 180
 ctgcatgggc caagattcca ttccttggaa tccgtgagac caagaacccc aggtcagaga 240
 acacaaggct tgccaccatg ttggaagcag cccaccacca ttttggaagc agcccgccac 300
 50 **tatcttyyya gctctgggag caaggacccc aggtaacaat ttggtgacca cgaagggacc 360**
 tgaatccgca accatgaagg gatctccaaa gcaattggaa atgttcctec caaggcaaaa 420
 atgcccctaa gatgtattct ggagaattgg gaccaatttg accctcagac agtaagaaaa 480
 55 **aaatgactta tattctcttg cagtaccgcc ctggccacga tatectcttc aagggggaga 540**
 aacctggcct cctgagggaa gtataaatta taacaccalc ttacagctag acctgttttg 600
 tagaaaagga ggcaaatgga gtgaagtgcc atatttacia actttctttt cattaanaaga 660
 caactcgcaa ttatgttaac agtgtgattt gtgttcctac acggaagccc tcagattcta 720
 60 **ctccccaccc ccggcatctc ccctgaatcc ctccccaaact tatt 764**

 <210> 142
 <211> 800
 65 <212> ADN
 <213> humano

