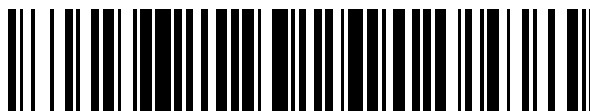


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 379 011**

51 Int. Cl.:
C12N 15/48 (2006.01)
C12N 7/00 (2006.01)
C07K 14/15 (2006.01)
C12Q 1/68 (2006.01)
C07K 16/10 (2006.01)
G01N 33/569 (2006.01)
A61K 39/21 (2006.01)
A61K 39/42 (2006.01)
A61K 48/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **07018564 .0**
96 Fecha de presentación: **02.08.1996**
97 Número de publicación de la solicitud: **1916304**
97 Fecha de publicación de la solicitud: **30.04.2008**

54 Título: **Polipéptidos antigénicos asociados a la esclerosis en placas y utilizaciones**

30 Prioridad:
03.08.1995 FR 9509643

45 Fecha de publicación de la mención BOPI:
19.04.2012

45 Fecha de la publicación del folleto de la patente:
19.04.2012

73 Titular/es:
BIO MERIEUX
69280 Marcy l'Etoile, FR

72 Inventor/es:
Perron, Hervé;
Beseme, Frédéric;
Bedin, Frédéric;
Paranhos-Baccala, Glaucia;
Komurian-Pradel, Florence;
Jolivet Reynaud, Colette y
Mandrand, Bernard

74 Agente/Representante:
Curell Aguilá, Mireia

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).

ES 2 379 011 T3

DESCRIPCIÓN

Polipéptidos antigénicos asociados a la esclerosis en placas y utilizaciones.

- 5 La esclerosis en placas (SEP) es una enfermedad desmielinizante del sistema nervioso central (SNC) cuya causa completa permanece todavía desconocida.

10 Numerosos estudios han apoyado la hipótesis sobre una etiología vírica de la enfermedad, pero ninguno de los virus conocidos estudiados se ha revelado como el agente causal buscado: se ha realizado un informe de los virus buscados desde hace años en la SEP por E. Norrby (1) y R. T. Jonson (2).

15 Recientemente, un retrovirus, diferente de los retrovirus humanos conocidos, ha sido aislado en pacientes afectados por la SEP (3, 4 y 5). Los autores han podido así demostrar que este retrovirus podía transmitirse *in vitro*, que pacientes afectados por la SEP producían anticuerpos susceptibles de reconocer proteínas asociadas a la infección de las células leptomeningeas por este retrovirus, y que la expresión de éste podía estar considerablemente estimulada por los genes inmediatos-precoces de ciertos herpesvirus (6).

20 Todos estos resultados apuntan hacia la implicación de al menos un retrovirus desconocido en la SEP o de un virus que tiene una actividad transcriptasa inversa detectable según el método publicado por H. Perron (3) y calificada de actividad "RT de tipo LM7".

25 Recientemente, los estudios del solicitante han permitido la obtención de dos líneas continuas de células infectadas por aislados naturales que proceden de pacientes diferentes afectados por la SEP, mediante un procedimiento de cultivo tal como se describe en el documento WO-A-93 20188. Estas dos líneas derivadas de células de plexos coroides humanos, denominadas LM7PC y PLI-2 han sido depositadas respectivamente en la E.C.A.C.C el 22 de julio de 1992 y el 8 de enero de 1993, con los números 92 072201 y 93 010817, según las disposiciones del Tratado de Budapest. Por otra parte, los aislados víricos que poseen una actividad RT de tipo LM7 han sido también depositados en la E.C.A.C.C. con la denominación global de "cepas". La "cepa" o aislado alojado por la línea PLI-2, denominada POL-2, ha sido depositada en la E.C.A.C.C el 22 de julio de 1992 con el nº V92072202. La "cepa" o aislado alojado por la línea LM7PC, denominada MS7PG, ha sido depositada en la E.C.A.C.C. el 8 de enero de 1993 con el nº V93010816.

35 A partir de los cultivos y de los aislados citados anteriormente, caracterizados por criterios biológicos y morfológicos, se ha involucrado en la caracterización del material nucleico asociado a las partículas virales producidas en estos cultivos.

40 Las porciones de genoma ya caracterizadas han sido utilizadas para elaborar ensayos de detección molecular del genoma vírico, y ensayos inmunoserológicos, utilizando las secuencias de aminoácidos codificadas por las secuencias de nucleótidos del genoma vírico, para detectar la respuesta inmunitaria dirigida contra epítomos asociados a la infección y/o a la expresión vírica.

45 Estas herramientas han permitido confirmar ya una asociación entre la SEP y la expresión de las secuencias identificadas en las patentes citadas a continuación. Sin embargo, el sistema vírico descubierto por el solicitante, es similar a un sistema retrovítico complejo. En efecto, las secuencias encontradas encapsidadas en las partículas víricas extracelulares producidas por los diferentes cultivos de células de pacientes afectados por la SEP, muestran claramente que existe una co-encapsidación de genomas retrovíricos similares, pero diferentes del genoma retrovítico "salvaje" que produce las partículas víricas infecciosos. Este fenómeno se ha observado entre retrovirus replicativos y retrovirus endógenos que pertenecen a la misma familia, incluso heterólogos. La noción de retrovirus endógeno es muy importante en el contexto del descubrimiento del solicitante porque, en el caso de MSRV-1, se ha observado que existen en el ADN humano normal unas secuencias retrovíricas endógenas que contienen secuencias homólogas al genoma MSRV-1. La existencia de elementos retrovíricos endógenos (ERV) similares a MSRV-1 en parte o en la totalidad de su genoma, explica el hecho de que la expresión del retrovirus MSRV-1 en las células humanas pueda interactuar con secuencias endógenas cercanas. Estas interacciones se vuelven a encontrar en el caso de retrovirus endógenos patógenos y/o infecciosos (por ejemplo algunas cepas ecotrópicas del virus de la leucemia murina), en el caso de retrovirus exógenos cuya secuencia nucleotídica puede encontrarse parcialmente o en su totalidad, en la 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 ERV por el retrovirus replicativo, (ii) una encapsidación "ilegítima" de ARN similares de ERV, o ERVS (incluso de ARN celulares) que posean simplemente secuencias de encapsidación compatibles, en las partículas retrovíricas producidas por la expresión de la cepa replicativa, a veces transmisibles y a veces con una patogenicidad propia, y (iii) recombinaciones más o menos importantes entre los genomas co-encapsidados, en particular en las fases de transcripción inversa, que llevan a la formación de genomas híbridos, a veces transmisibles y a veces con una patogenicidad propia.

65 De esta manera, (i) se han encontrado diferentes secuencias relacionadas con MSRV-1 en las partículas víricas purificadas; (ii) el análisis molecular de diferentes regiones del genoma retrovítico MSRV-1 se debe realizar

analizando sistemáticamente las secuencias co-encapsidadas, interferentes y/o recombinadas generadas por la infección y/o la expresión de MSRV-1, además, ciertos clones pueden tener partes de secuencias defectivas producidas por la replicación retroviral y los errores de matriz y/o de transcripción de la transcriptasa inversa; (iii) las familias de secuencias relacionadas con una misma región geonómica retroviral son los soportes de una detección diagnóstica global que se puede optimizar mediante la identificación de regiones invariables entre los clones expresados y mediante la identificación de tramas de lecturas responsables de la producción de polipéptidos antigénicos y/o patógenos que se pueden producir únicamente mediante una parte, incluso por uno solo, de los clones expresados y, en estas condiciones, el análisis sistemático de los clones expresados en una región de un genoma dado permite evaluar la frecuencia de variación y/o de recombinación del genoma MSRV-1 en esta región y definir las secuencias óptimas para las aplicaciones, en particular diagnósticas; (iv) la patología provocada por un retrovirus como el MSRV-1 puede ser un efecto directo de su expresión y de las proteínas o de los péptidos producidos por este hecho, pero también un efecto de la activación, de la encapsidación, de la recombinación de los genomas relacionados o heterólogos, y de las proteínas o de los péptidos producidos por estos hechos; de esta manera estos genomas asociados a la expresión y/o la infección por MSRV-1 son una parte integrante de la patogenidad potencial de este virus y, por lo tanto, constituyen soportes de detección diagnóstica y dianas terapéuticas concretas. De la misma manera, cualquier agente que esté asociado a un cofactor o sea un cofactor de estas interacciones responsables de la patogenia en cuestión, tal como MSRV-2 o el factor gliotóxico descritos en la solicitud de patente publicada con el nº 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 descubierto paralelamente en otra enfermedad autoinmune, la poliartritis reumatoide (PR), que se describió en la solicitud de patente francesa depositada con el nº 95 02960. Este descubrimiento demuestra que, aplicando enfoques metodológicos similares a los que fueron utilizados en los estudios del solicitante sobre la SEP, se ha podido identificar un retrovirus expresado en la PR que comparte las secuencias descritas para MSRV-1 en la SEP, y también la coexistencia de una secuencia asociada a MSRV-2 también descrita en la SEP. En lo que se refiere a MSRV-1, las secuencias detectadas de manera común a la SEP y la 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 MSRV-1 expresadas en estas dos enfermedades.

La presente solicitud de patente tiene por objeto diferentes resultados, que se suman a los que ya están protegidos por las solicitudes de patente francesas:

- nº 92 04322 del 03.04.1992, publicada con el nº 2 689 519;
- nº 92 13447 del 03.11.1992, publicada con el nº 2 689 521;
- nº 92 13443 del 03.11.1992, publicada con el nº 2 689 520;
- nº 94 01529 del 04.02.1994, publicada con el nº 2 715 936;
- nº 94 01531 del 04.02.1994, publicada con el nº 2 715 939;
- nº 94 01530 del 04.02.1994, publicada con el nº 2 715 936;
- nº 94 01532 del 04.02.1994, publicada con el nº 2 715 937;
- nº 94 14322 del 24.11.1994, publicada con el nº 2 727 428; y
- nº 94 15810 del 23.12.1994, publicada con el nº 2 728 585;

La presente invención es una solicitud divisionaria de la solicitud EP 96420265.9 (EP 789 077) en la que el solicitante ha determinado un material vírico, en estado aislado o purificado, que puede ser aprehendido o estar caracterizado de diferentes maneras:

- su genoma comprende una secuencia nucleotídica seleccionada de entre el grupo constituido por las secuencias SEC ID nº 46, SEC ID nº 51, SEC ID nº 52, SEC ID nº 53, SEC ID nº 56, SEC ID nº 58, SEC ID nº 59, SEC ID nº 60, SEC ID nº 61, SEC ID nº 89, presentando sus secuencias complementarias, y sus secuencias equivalentes, en particular las secuencias nucleotídicas, para cualquier serie de 100 monómeros contiguos, por lo menos 50% y preferentemente por lo menos 70% de homología con respectivamente dichas secuencias SEC ID nº 46, SEC ID nº 51, SEC ID nº 52, SEC ID nº 53, SEC ID nº 56, SEC ID nº 58, SEC ID nº 59, SEC ID nº 60, SEC ID nº 61, SEC ID nº 89, y sus secuencias complementarias;
- la región de su genoma que comprende los genes *env*, *pol* y una parte de gen *gag*, excepto la sub-región que tiene una secuencia idéntica o equivalente a la SEC ID nº 1, codifica para cualquier polipéptido que presenta, para cualquier serie contigua de por lo menos 30 aminoácidos, por lo menos 50%, y preferentemente por lo menos 70% de homología con una secuencia peptídica codificada por cualquier secuencia nucleotídica seleccionada de entre el grupo constituido por SEC ID nº 46, SEC ID nº 51, SEC ID nº 52, SEC ID nº 53, SEC ID nº 56, SEC ID nº 58, SEC ID nº 59, SEC ID nº 60, SEC ID nº 61, SEC ID nº 89, y sus secuencias complementarias;
- el gen *pol* comprende una secuencia nucleotídica idéntica o equivalente, parcial o totalmente, a la SEC ID nº 57, excepto la SEC ID nº 1;

- el gen *gag* comprende una secuencia nucleotídica idéntica o equivalente, parcial o totalmente, a la SEC ID nº 88.

Tal como se ha mencionado anteriormente, el material vírico tal como se ha definido anteriormente está asociado a la SEP. Y tal como se ha definido haciendo referencia al gen *pol* o *gag* de MSRV-1, y más particularmente a las

El solicitante ha definido asimismo diferentes fragmentos nucleotídicos, que comprenden cada uno, una secuencia nucleotídica seleccionada de entre el grupo constituido por:

- (a) todas las secuencias genómicas, parciales y totales, del gen *pol* del virus MSRV-1, excepto la secuencia total del fragmento nucleotídico definido por SEC ID nº 1;

- (b) todas las secuencias genómicas, parciales y totales, del gen *env* de MSRV-1;

- (c) todas las secuencias genómicas parciales del gen *gag* de MSRV-1;

- (d) todas las secuencias genómicas que solapan el gen *pol* y el gen *env* del virus MSRV-1, y que solapan el gen *pol* y el gen *gag*;

- (e) todas las secuencias, parciales y totales, de un clon seleccionado de entre el grupo constituido por los clones FBd3 (SEC ID nº 46), t *pol* (SEC ID nº 51), JLBc1 (SEC ID nº 52), JLBc2 (SEC ID nº 53), GM3 (SEC ID nº 56), FBd13 (SEC ID nº 58), LB19 (SEC ID nº 59), LTRGAG12 (SEC ID nº 60), FP6 (SEC ID nº 61), G+E+A (SEC ID nº 89), excepto cualquier secuencia nucleotídica idéntica a, o comprendida en, la secuencia definida por SEC ID nº 1;

- (f) las secuencias complementarias a dichas secuencias genómicas;

- (g) las secuencias equivalentes a dichas secuencias (a) a (e), en particular las secuencias nucleotídicas que presentan, para cada serie de 100 monómeros contiguos, por lo menos 50% y preferentemente por lo menos 70% de homología con dichas secuencias (a) a (d),

con la condición de que este fragmento nucleotídico no comprenda o no consista en la secuencia ERV-9 tal como se describe en LA MANTIA *et al* (18).

- Mediante secuencias genómicas, parciales o totales, se incluyen todas las secuencias asociadas mediante coencapsidación o mediante coexpresión, o recombinadas.

Preferentemente, dicho fragmento comprende:

- o una secuencia nucleotídica idéntica a una secuencia genómica parcial o total del gen *pol* del virus MSRV-1, excepto a la secuencia total del fragmento nucleotídico definido por SEC ID nº 1, o idéntica a cualquier secuencia equivalente a dicha secuencia genómica parcial o total, en particular homóloga a esta última;

- o una secuencia nucleotídica idéntica a una secuencia genómica parcial o total del gen *env* del virus MSRV-1, o idéntica a cualquier secuencia complementaria de dicha secuencia nucleotídica, o idéntica a cualquier secuencia equivalente a dicha secuencia nucleotídica, en particular homóloga a esta última.

Según el documento H. Perron *et al.*, Research in Virology, vol. 143 N°5 (1992) p337-350, los autores han analizado el perfil antigénico del retrovirus LM7 aislado de células leptomenígeas de un paciente afectado de SEP. Mediante técnicas de transferencia Western y de inmunoprecipitación, han puesto en evidencia la existencia de antígenos víricos que han caracterizado por su peso molecular.

Se describen diferentes péptidos codificados por cualquier marco abierto de lectura cuya secuencia nucleotídica pertenece al gen *pol* del virus MSRV-1 tal como el definido anteriormente, y en particular cualquier polipéptido antigénico reconocido por sueros de pacientes infectados por el virus MSRV-1, y/o en los cuales se ha reactivado el virus MSRV-1.

A título de ejemplo, dicho polipéptido está codificado por cualquier marco abierto de lectura cuya secuencia

nucleotídica pertenece a la SEC ID nº 1, SEC ID nº 89 o a su secuencia complementaria. Así, está codificado ventajosamente por el marco abierto de lectura que empieza en el sentido 5'-3', en el nucleótido 181, y que termina en el nucleótido 330 de la SEC ID nº 1 o el marco abierto de lectura que empieza en el sentido 5'-3', en el nucleótido 529, y que termina en el nucleótido 606 de la SEC ID nº 1 o también el marco abierto de lectura que empieza en el sentido 5'-3', en el nucleótido 1118, y que termina en el nucleótido 1267 de SEC ID nº 89, o también el marco abierto de lectura que empieza en el sentido 5'-3', en el nucleótido 1464, y que termina en el nucleótido 1541 de SEC ID nº 89.

La presente invención tiene por objeto un polipéptido antigénico, reconocido por los sueros de pacientes infectados por el virus MSRV-1, y/o en los que se ha reactivado el virus MSRV-1, cuya secuencia peptídica consiste en una secuencia seleccionada de entre SEC ID nº 39, SEC ID nº 41 a SEC ID nº 44, SEC ID nº 63.

Un polipéptido preferido asimismo posee una secuencia peptídica que consiste en una secuencia seleccionada de entre SEC ID nº 39 y SEC ID nº 63.

La invención se refiere asimismo a una secuencia nucleotídica que codifica para un polipéptido tal como el definido anteriormente, seleccionada de entre SEC ID nº 40, SEC ID nº 62, sus fragmentos y sus secuencias complementarias de éstos.

La presente invención propone asimismo anticuerpos mono- o policlonales, dirigidos contra el virus MSRV-1, obtenidos mediante reacción inmunológica de un organismo humano o animal, con un agente inmunógeno constituido por un polipéptido antigénico tal como el definido anteriormente.

La invención se refiere asimismo:

- a los agentes reactivos de detección del virus MSRV-1, o de una exposición a este último, que comprenden a título de sustancia reactiva, un polipéptido, en particular antigénico, tal como se ha definido anteriormente, o un anticuerpo de dicho polipéptido;

- a cualquier composición diagnóstica, profiláctica o terapéutica, que comprende uno o varios polipéptidos, tales como se han definido anteriormente, o uno o varios anticuerpos de los polipéptidos considerados anteriormente; preferentemente, dicha composición es, a título de ejemplo, una composición inmunoterapéutica activa, en particular una composición vacunal.

La invención se refiere asimismo a cualquier composición diagnóstica, profiláctica o terapéutica, en particular para inhibir la expresión de al menos un agente patógeno y/o infeccioso asociado a la SEP, que comprende una secuencia nucleotídica tal como la definida anteriormente. De la misma manera, se refiere a cualquier composición diagnóstica, profiláctica o terapéutica, en particular para inhibir la expresión de al menos un agente patógeno y/o infeccioso asociado a la PR, que comprende un fragmento nucleotídico tal como se ha definido anteriormente con respecto a los genes pol, y en particular con respecto a las secuencias SEC ID nº 40, 62 y 89.

Estos mismos fragmentos, o polinucleótidos, en particular oligonucleótidos, pueden entrar en cualquier composición apropiada, para detectar, según cualquier procedimiento o método conveniente, un agente patológico y/o infeccioso, asociado respectivamente a la SEP y a la PR, en una muestra biológica. En dicho procedimiento, se pone en contacto un ARN y/o un ADN que pertenecen o proceden supuestamente de dicho agente patológico y/o infeccioso, y/o su ARN y/o ADN complementario, con dicha composición.

La presente invención se refiere asimismo a cualquier procedimiento para detectar la presencia de dicho agente o la exposición a dicho agente patológico y/o infeccioso, en una muestra biológica, poniendo en contacto esta muestra con un polipéptido, tal como se ha definido anteriormente, o anticuerpo de este polipéptido, tal como se ha definido anteriormente.

En la práctica, y por ejemplo, un dispositivo de detección del virus MSRV-1 comprende un agente reactivo tal como se ha definido anteriormente, soportado por un soporte sólido, inmunológicamente compatible con el agente reactivo, y un medio de puesta en contacto de la muestra biológica, por ejemplo una muestra de sangre o de líquido cefalorraquídeo, susceptible de contener anticuerpos anti-MSRV-1, con este agente reactivo, en condiciones que permiten una eventual reacción inmunológica, con medios de detección del complejo inmune formado con este agente reactivo.

La invención se refiere asimismo a la detección de anticuerpos anti-MSRV-1, en una muestra biológica, por ejemplo una muestra de sangre o de líquido cefalorraquídeo, según la cual se pone en contacto esta muestra con un agente reactivo tal como se ha definido anteriormente, que consiste en un anticuerpo, en condiciones que permiten su eventual reacción inmunológica y, después, se detecta la presencia del complejo inmune así formado con el agente reactivo.

Antes de describir con mayor detalle la invención, se definirán a continuación diferentes términos utilizados en la descripción y en las reivindicaciones.

- 5 - por cepa o aislado, se entiende cualquier fracción biológica infecciosa y/o patógena, que contiene por ejemplo virus y/o bacterias y/o parásitos, que genera un poder patógeno y/o antigénico, acogida por un cultivo o un hospedante vivo; a título de ejemplo, una cepa vírica según la definición anterior puede contener un agente coinfeccioso, por ejemplo un protista patógeno,
- 10 - el término "MSRV" utilizado en la presente descripción se refiere a cualquier agente patógeno y/o infeccioso, asociado a la SEP, en particular una especie vírica, las cepas atenuadas de dicha especie vírica, o las partículas defectivas interferentes o que contienen genomas coencapsidados o también genomas recombinados con una parte del genoma MSRV-1, derivadas de esta especie. Se sabe que los virus y particularmente los virus que contienen ARN tienen una variabilidad consecutiva en particular a índices relativamente elevados de mutación espontánea (7), la cual se tendrá en cuenta más adelante para definir la noción de equivalencia,
- 15 - mediante la expresión "virus humano", se entiende un virus susceptible de infectar o de ser acogido por el ser humano,
- 20 - teniendo en cuenta todas las variaciones y/o las 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,
- 25 - la variante de un virus o de un agente patógeno y/o infeccioso, según la invención, comprende al menos un antígeno reconocido por al menos un anticuerpo dirigido contra al menos un antígeno correspondiente de dicho virus y/o de dicho agente patógeno y/o infeccioso, y/o un genoma del que se detecta cualquier parte mediante al menos una sonda de hibridación, y/o al menos un cebador de amplificación nucleotídico específico de dicho virus y/o agente patógeno y/o infeccioso, como por ejemplo para el virus denominado MSRV-1, las que tienen una secuencia nucleotídica seleccionada de entre SEC ID nº 20 a SEC ID nº 24, SEC ID nº 26, SEC ID nº 16 a SEC ID nº 19, SEC ID nº 31 a SEC ID nº 33, SEC ID nº 45, SEC ID nº 47, SEC ID nº 48, SEC ID nº 49, SEC ID nº 50, SEC ID nº 45 y sus secuencias complementarias, en condiciones de hibridación determinadas bien conocidas por el experto en la materia,
- 30 - 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 condiciones predeterminadas, pudiendo la cadena contener monómeros de estructuras químicas diferentes y ser obtenida a partir de una molécula de ácido nucleico natural y/o mediante recombinación genética y/o mediante 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;
- 35 - asimismo 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 se trate del ADN o del ARN, la base nitrogenada se selecciona de entre la adenina, la guanina, el uracilo, la citosina, la timina; o el nucleótido se puede modificar por lo menos en uno de los tres elementos constitutivos; a título de ejemplo, la modificación puede intervenir al nivel de las bases, generando bases modificadas tales como la inosina, la metil-5-desoxicidina, la desoxiuridina, la dimetilamino-5-desoxiuridina, la diamino-2,6-purina, la bromo-5-desoxiuridina y cualquier otra base modificada que favorece la hibridación; al nivel del azúcar, la modificación puede consistir en la sustitución de al menos una desoxirribosa por una poliamida (8), y al nivel del grupo fosfato, la modificación puede consistir en su sustitución por ésteres, específicamente seleccionados de entre los ésteres de difosfato, de alquilo y arilfosfonato y de fosforotioato,
- 40 - mediante la expresión "secuencia informacional", se entiende cualquier serie ordenada de monómeros, cuya naturaleza química y cuyo orden en un sentido de referencia, constituyen o no una información funcional de la misma calidad que la de los ácidos nucleicos naturales,
- 45 - por hibridación, se entiende el proceso durante el cual, en condiciones operativas apropiadas, dos fragmentos nucleotídicos, que tienen secuencias suficientemente complementarias, se emparentan para formar una estructura compleja, en particular doble o triple, preferentemente en forma de hélice,
- 50 - una sonda contiene un fragmento nucleotídico sintetizado por vía química u obtenido por digestión o corte enzimático de un fragmento nucleotídico más largo, que comprende al menos seis monómeros, ventajosamente entre 10 y 100 monómeros, preferentemente entre 10 y 30 monómeros, y que posee una especificidad de hibridación en condiciones determinadas; preferentemente, una sonda que posee menos de 10 monómeros no se utiliza sola, pero sí en presencia de otras sondas de tamaño tan corto o no; en ciertas condiciones particulares, puede resultar útil utilizar sondas de tamaño superior a 100 monómeros; en particular, se puede
- 55
- 60
- 65

usar una sonda con fines de diagnóstico, y se tratará por ejemplo de sondas de captura y/o de detección,

- 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 absorción pasiva,
- la sonda de detección se puede marcar mediante un marcador seleccionado en particular de entre los isótopos radioactivos, enzimas seleccionadas en particular de entre la peroxidasa y la fosfatasa alcalina, y los susceptibles de hidrolizar un sustrato cromógeno, fluorígeno o luminiscente, compuestos químicos cromóforos, compuestos cromógenos, fluorígenos o luminiscentes, análogos de bases nucleotídicas, y la biotina,
- en el marco de la presente invención, las sondas utilizadas con fines de diagnóstico se pueden implementar en cualquier técnica de hibridación conocida, y en particular las técnicas denominadas "TRANSFERENCIA DE DOT" (9), "TRANSFERENCIA SOUTHERN" (10), "TRANSFERENCIA NORTHERN" que es una técnica idéntica a la técnica de "TRANSFERENCIA SOUTHERN" pero que usa ARN como diana, la técnica "SÁNDWICH" (11); ventajosamente, en el marco de la presente invención se usa la técnica de "SÁNDWICH", 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 al menos parcialmente diferente,
- en el marco de la presente invención, cualquier sonda puede hibridarse *in vivo* o *in vitro* sobre el ARN y/o el ADN, para bloquear los fenómenos de replicación, en particular de traducción y/o de transcripción, y/o para degradar dicho ADN y/o ARN,
- un cebador es una sonda que comprende al menos seis monómeros, y ventajosamente de 10 a 30 monómeros, que poseen una especificidad de hibridación en condiciones determinadas, para la iniciación de una polimerización enzimática, por ejemplo en una técnica de amplificación tal como PCR (Polymerase Chain Reaction), en un procedimiento de elongación, tal como la secuenciación, en un método de transcripción inversa o análogo,
- dos secuencias nucleotídicas o peptídicas se denominan equivalentes o derivadas una con relación a la otra, o con relación a una secuencia de referencia, si de manera funcional los biopolímeros correspondientes pueden desempeñar sustancialmente la misma función, sin ser idénticas, en relación con la aplicación o la utilización considerada, o en la técnica en la que intervienen; son en particular equivalentes dos secuencias obtenidas por el hecho de la variabilidad natural, en particular mutación espontánea de la especie a partir de la cual se identifican, o inducida, así como dos secuencias homólogas, definiéndose la homología a continuación,
- por "variabilidad", se entiende cualquier modificación, espontánea o inducida de una secuencia, en particular por sustitución, y/o inserción, y/o supresión de nucleótidos y/o de fragmentos nucleotídicos, y/o extensión y/o acortamiento de la secuencia en por lo menos uno de los extremos; una variabilidad no natural puede resultar de las técnicas de ingeniería genética utilizadas, por ejemplo de la elección de los cebadores de síntesis degenerados o no, considerados para amplificar un ácido nucleico; esta variabilidad se puede traducir por modificaciones de cualquier secuencia de partida, considerada como referencia, y pudiendo ser expresadas por un grado de homología en relación con dicha secuencia de referencia,
- la homología caracteriza el grado de identidad de dos fragmentos nucleotídicos o peptídicos comparados; se mide por el porcentaje de identidad que es específicamente determinado por comparación directa de secuencias nucleotídicas o peptídicas, en relación con secuencias nucleotídicas o peptídicas de referencia,
- este porcentaje de identidad se ha determinado específicamente para los fragmentos nucleotídicos, en particular clones mencionados anteriormente, homólogos a los fragmentos identificados para el virus MSRV-1 por las SEC ID nº 1 a nº 9, SEC ID nº 46, SEC ID nº 51 a SEC ID nº 53, SEC ID nº 40, SEC ID nº 56 y SEC ID nº 57, así como para las sondas y cebadores homólogos a las sondas y cebadores identificados por las SEC ID nº 20 a SEC ID nº 24, SEC ID nº 26, SEC ID nº 16 a SEC ID nº 19, SEC ID nº 31 a SEC ID nº 33, SEC ID nº 45, SEC ID nº 47, SEC ID nº 48, SEC ID nº 49, SEC ID nº 50, SEC ID nº 55, SEC ID nº 40, SEC ID nº 56 y SEC ID nº 57; a título de ejemplo, el porcentaje más bajo de identidad observado entre los diferentes consensos generales de ácidos nucleicos obtenidos a partir de fragmentos de ARN vírico de MSRV-1, procedente de las líneas LM7PC y PLI-2 según un protocolo detallado más adelante, es de 67% en la región descrita en la figura 1,
- 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, son en particular equivalentes a un fragmento de nucleótidos de referencia:
 - (a) cualquier fragmento susceptible de hibridarse al menos parcialmente con el complemento del fragmento de referencia,
 - (b) cualquier fragmento cuya alineación con el fragmento de referencia conduce a poner en evidencia 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,

(d) cualquier fragmento que puede resultar de técnicas de ingeniería genética aplicadas al fragmento de referencia,

(e) cualquier fragmento, que contiene al menos ocho nucleótidos contiguos, que codifica para un péptido homólogo o idéntico al péptido codificado por el fragmento de referencia,

(f) cualquier fragmento diferente del fragmento de referencia, por inserción, supresión, sustitución de al menos un monómero, extensión, o acortamiento de al menos uno de sus extremos; por ejemplo, cualquier fragmento que corresponde al fragmento de referencia, flanqueado en al menos uno de sus extremos por una secuencia nucleotídica que no codifica para un polipéptido,

- mediante el término "polipéptido", se entiende en particular cualquier péptido de al menos dos aminoácidos, en particular oligopéptido, proteína, extracto, separado, o sustancialmente aislado o sintetizado, por la intervención del ser humano, en particular los que se obtienen mediante síntesis química, o mediante expresión en un organismo recombinante,

- mediante la expresión "polipéptido codificado de manera parcial por un fragmento nucleotídico", se entiende un polipéptido que presenta al menos tres aminoácidos codificados por al menos nueve monómeros contiguos comprendidos en dicho fragmento nucleotídico,

- 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.

- 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; es en particular equivalente a un polipéptido de referencia:

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

(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,

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

(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 aminas, una carboxilación de las funciones tiol, una esterificación de las funciones carboxílicas,

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

(g) cualquier polipéptido del que al 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 de al menos 50% y preferentemente de al menos 70%.

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

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

Mediante la expresión "detección de una sustancia o de un 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, que hace referencia a las figuras anexas, en las que:

- la figura 1 representa unos consensos generales de ácidos nucleicos de los clones MSRV-1B amplificados mediante la técnica PCR en la región "pol" definida por Shih (12), a partir de ADN vírico que procede de las líneas LM7PC y PLI-2, e identificados con las referencias SEC ID nº 3, SEC ID nº 4, SEC ID nº 5 y SEC ID nº 6, y el consenso común con cebadores de amplificación que tienen la referencia SEC ID nº 7.
- la figura 2 da la definición de una trama de lectura funcional para cada familia de tipo MSRV-1B/"PCR pol", siendo dichas familias A a D definidas respectivamente por las secuencias nucleotídicas SEC ID nº 3, SEC ID nº 4, SEC ID nº 5 y SEC ID nº 6, descritas en la figura 1.
- la figura 3 da un ejemplo de consenso de las secuencias de MSRV-2B, identificado mediante la SEC ID nº 11,
- la figura 4 es una representación de la actividad de transcriptasa inversa (RT) en dpm (desintegración por minuto), en las fracciones de sacarosa extraídas en un gradiente de purificación de los viriones producidos por los linfocitos B en cultivo de un paciente que padece SEP,
- la figura 5 da, en las mismas condiciones experimentales que en la figura 4, la determinación de la actividad de transcriptasa inversa en el cultivo de una línea de linfocitos B obtenida a partir de un control libre de SEP,
- la figura 6 representa la secuencia nucleotídica del clon PSJ17 (SEC ID nº 9),
- la figura 7 representa la secuencia nucleotídica SEC ID nº 8, del clon denominado M003-P004,
- la figura 8 representa la secuencia nucleotídica SEC ID nº 2 del clon F11-1; la parte observada entre las dos flechas en la región del cebador corresponde a una variabilidad impuesta por la elección del cebador que ha servido a la clonación de F11-1; en esta misma figura, se representa la traducción en aminoácidos,
- la figura 9, que agrupa las figuras 9a y 9b, representa la secuencia nucleotídica SEC ID nº 1, y un tramo funcional de lectura posible de aminoácidos de SEC ID nº 1; en esta secuencia, las secuencias de consenso del gen pol están subrayadas,
- las figuras 10 y 11 dan los resultados de una PCR, en forma de fotografía bajo luz ultravioleta de un gel de azarosa impregnado de bromuro de etidio, de los productos obtenidos a partir de los cebadores identificados mediante la SEC ID nº 16, SEC ID nº 17, SEC ID nº 18 y SEC ID nº 19.
- la figura 12 muestra una representación matricial de la homología entre la SEC ID nº 1 de MSRV-1 y la de un retrovirus endógeno denominado HSERV9; esta homología de al menos 65% se muestra mediante un trazo lleno, significando la ausencia de trazo una homología inferior a 65%,
- la figura 13 representa la secuencia nucleotídica SEC ID nº 46 del clon FBd3,
- la figura 14 representa la homología de secuencia entre el clon FBd3 y el retrovirus HSERV-9;
- la figura 15 representa la secuencia nucleotídica SEC ID nº 51 del clon t pol,
- las figuras 16 y 17 representan respectivamente las secuencias nucleotídicas SEC ID nº 52 y SEC ID nº 53, de los clones JLBc1 y JLBc2 respectivamente;
- la figura 18 representa la homología de secuencia entre el clon JLBc1 y el clon FBd3,
- la figura 19 representa la homología de secuencia entre el clon JLBc2 y el clon FBd3,
- la figura 20 representa la homología entre los clones JLBc1 y JLBc2,
- las figuras 21 y 22 representan la homología de secuencia entre el retrovirus HSERV-9 y los clones JLBc1 y JLBc2 respectivamente,
- la figura 23 representa la secuencia nucleotídica SEC ID nº 56 del clon GM3;
- la figura 24 representa la homología de secuencia entre el retrovirus HSERV-9 y el clon GM3,

- la figura 25 representa la localización de los diferentes clones estudiados, con relación al genoma del retrovirus conocido ERV9,
- 5 - la figura 26 representa la posición de los clones F11-1, M003-P004, MSRV-1B y PSJ17 en la región denominada a continuación MSRV-1 pol*,
- la figura 27, explosionada en tres figuras sucesivas 27a, 27b, 27c, representa un cuadro de lectura posible, que cubre la totalidad del gen pol,
- 10 - la figura 28 representa según la SEC ID nº 40, la secuencia nucleotídica que codifica para el fragmento peptídico POL2B, que tiene la secuencia de aminoácidos identificada mediante la SEC ID nº 39,
- la figura 29 representa los valores de DO (ensayos ELISA) a 492 nm obtenidos para 29 sueros de pacientes SEP y 32 sueros de pacientes de control sanos ensayados con un anticuerpo anti-IgG,
- 15 - la figura 30 representa los valores de DO (ensayos ELISA) a 492 nm obtenidos para 36 sueros de pacientes SEP y 42 sueros de pacientes de control sanos ensayados con un anticuerpo anti-IgM,
- las figuras 31 a 33 representan los resultados obtenidos (intensidad relativa de los puntos) para 43 octapéptidos solapantes que cubren la secuencia de aminoácidos 61-110, según la técnica de SpotsScan, respectivamente con un grupo de sueros de SEP, con un grupo de sueros de control y con el grupo de sueros SEP después de la deducción de un ruido de fondo que corresponde a la señal máxima detectada en al menos un octapéptido con el suero de control (intensidad = 1), entendiéndose que estos sueros han sido diluidos al 1/50. La barra en el extremo derecho representa un estándar de escala gráfica sin relación con el ensayo serológico,
- 20 - la figura 34 representa las SEC ID nº 41 y SEC ID nº 42 de dos polipéptidos que comprenden inmunodominantes, mientras que las SEC ID nº 43 y 44 representan polipéptidos inmunorreactivos y específicos de SEP,
- 25 - la figura 35 representa la secuencia nucleotídica SEC ID nº 59 del clon LB19 y tres tramos de lectura potenciales de aminoácidos de SEC ID nº 59,
- la figura 36 representa la secuencia nucleotídica SEC ID nº 88 (GAG*) y una trama de lectura potencial de aminoácidos de SEC ID nº 88,
- 35 - la figura 37 representa la homología de secuencia entre el clon FBd13 y el retrovirus HSERV-9; según esta representación, el trazo lleno significa un porcentaje de homología superior o igual a 70%, y la ausencia de trazo significa un porcentaje de homología inferior,
- 40 - la figura 38 representa la secuencia nucleotídica SEC ID nº 61 del clon FP6 y tres tramas de lectura potenciales de aminoácidos de SEC ID nº 61,
- la figura 39 representa la secuencia nucleotídica SEC ID nº 89 del clon G+E+A y tres tramas de lectura potenciales de aminoácidos de SEC ID nº 89,
- 45 - la figura 40 representa una trama de lectura encontrada en la región E y que codifica para una proteasa retrovítica MSRV-1 identificada por la SEC ID nº 90;
- la figura 41 representa la respuesta de cada suero de pacientes que padecen SEP indicada mediante el símbolo (+) y de pacientes sanos simbolizada mediante (-), ensayada con un anticuerpo anti-IgG expresada en densidad óptica neta a 492 nm, y
- 50 - la figura 42 representa la respuesta de cada suero de pacientes que padecen SEP indicado mediante los símbolos (+) y (QS) y de pacientes sanos (-), ensayado con un anticuerpo anti-IgM expresada en densidad óptica neta a 492 nm.
- 55

EJEMPLO 1: OBTENCIÓN DE CLONES DENOMINADOS MSRV-1B Y MSRV-2B, QUE DEFINICEN RESPECTIVAMENTE UN RETROVIRUS MSRV-1 Y UN AGENTE COINFECCIOSO MSRV2, MEDIANTE AMPLIFICACIÓN PCR "ANIDADA" DE LAS REGIONES POL CONSERVADAS DE LOS RETROVIRUS, EN PREPARACIONES DE VIRIONES QUE PROCEDEN DE LAS LÍNEAS LM7PC Y PLI-2

Se ha usado una técnica de PCR derivada de la técnica publicada por Shih (12). Esta técnica permite, mediante tratamiento de todos los componentes del medio de reacción por la ADNasa, eliminar cualquier traza de ADN contaminante. Permite paralelamente, mediante el uso de cebadores diferentes pero solapantes en dos series sucesivas de ciclos de amplificaciones PCR, aumentar las posibilidades de amplificar un ADNc sintetizado a partir de una cantidad de ARN baja al principio, y también reducida en la muestra mediante la acción parásita de la ADNasa

sobre el ARN. En efecto, la ADNasa se usa en condiciones de actividad en exceso que permiten eliminar cualquier traza de ADN contaminante, antes de la desactivación de esta enzima que se queda en la muestra mediante calentamiento hasta 85°C durante 10 minutos. Esta variante de la técnica de PCR descrita por Shih (12), se ha usado sobre un ADNc sintetizado a partir de los ácidos nucleicos de fracciones de partículas infecciosas purificadas sobre gradiente de sacarosa, según la técnica descrita por H-Perron (13), a partir del aislado "POL-2" (ECACC n°V92072202) producido por la línea PLI-2 (ECACC n°92072201) por una parte, y a partir del aislado MS7PG (ECACC n°V93010816) producido por la línea LM7PC (ECACC n°93010817) por otra parte. Estos cultivos se han obtenido según los métodos que han sido objeto de las solicitudes de patentes publicadas con los n° WO 93/20188 y WO 93/20189.

Después de la clonación con TA Cloning kit® de los productos amplificados mediante esta técnica, y análisis de la secuencia con la ayuda de un secuenciador automático "Automatic Sequencer, modelo 373A" de Applied Biosystems, se analizaron las secuencias con la ayuda del programa Beneworks®, en la última versión disponible del banco de datos Genbank®.

Las secuencias clonadas y secuenciadas a partir de estas muestras corresponden en particular a dos tipos de secuencias: un primer tipo de secuencia, encontrado en la mayoría de los clones (55% de los clones que proceden de los aislados POL-2 de los cultivos PLI-2, y 67% de los clones que proceden de los aislados MS7PG de los cultivos LM7PC), que corresponde a una familia de secuencias "pol" cercanas pero diferentes del retrovirus endógeno humano denominado ERV-9 o HSERV-9, y un segundo tipo de secuencia que corresponde a secuencias muy homólogas a una secuencia atribuida a otro agente infeccioso y/o patógeno denominado MSRV-2.

El primer tipo de secuencias que representa la mayoría de los clones está constituido por secuencias cuya variabilidad permite definir cuatro subfamilias de secuencias. Estas subfamilias son suficientemente cercanas entre sí para que se las pueda considerar como casi-especies que proceden de un mismo retrovirus, tal como se conoce bien para el retrovirus VIH-1 (14), o como el resultado de la interferencia con varios provirus endógenos co-regulados en las células productoras. Estos elementos endógenos más o menos defectivos son sensibles a las mismas señales de regulación generadas eventualmente por un provirus replicativo, debido a que pertenecen a la misma familia de retrovirus endógenos (15). Esta nueva familia de retrovirus endógenos o, alternativamente, esta nueva especie retroviral de la que se obtuvo en cultivo la generación de casi-especies, y que contiene un consenso de las secuencias descritas anteriormente, se denomina MSRV-1B.

En la figura 1 se representan los consensos generales de las secuencias de los diferentes clones MSRV-1B secuenciados durante este experimento, siendo estas secuencias respectivamente identificadas mediante la SEC ID n° 3, SEC ID n° 4, SEC ID n° 5 y SEC ID n° 6. Estas secuencias presentan una homología de ácidos nucleicos comprendida entre 70% y 88% con la secuencia HSERV9 referenciada X57147 y M37638 en la base de datos Genbank®. Se han definido cuatro secuencias nucleicas "consenso" representativas de diferentes casi-especies de un retrovirus MSRV-1B eventualmente exógeno, o de diferentes subfamilias de un retrovirus endógeno MSRV-1b. Estos consensos representativos se representan en la figura 2, con la traducción de aminoácidos. Una trama de lectura funcional existe para cada subfamilia de estas secuencias MSRV-1B, y se puede observar que la trama de lectura abierta funcional corresponde cada vez a la secuencia de aminoácidos que vienen en segunda línea bajo la secuencia de ácidos nucleico. El consenso general de la secuencia MSRV-1B, identificado mediante la SEC ID n° 7 y obtenido mediante esta técnica de PCR en la región "pol" se representa en la figura 1.

El segundo tipo de secuencia que representa la mayoría de los clones secuenciados se representa mediante la secuencia MSRV-2B representada en la figura 3 e identificada mediante la SEC ID n° 11. Las diferencias observadas en las secuencias que corresponden a los cebadores de PCR se explican mediante el uso de cebadores degenerados en mezcla usados en condiciones técnicas diferentes.

La secuencia MSRV-2B (SEC ID n° 11) es suficientemente divergente de las secuencias retrovirales ya descritas en los bancos de datos, para que se pueda anticipar que se trata de una región de secuencia que pertenece a un nuevo agente infeccioso, denominado MSRV-2. Este agente infeccioso se emparentaría *a priori*, según el análisis de las primeras secuencias obtenidas, a un retrovirus, pero, debido a la técnica usada para obtener esta secuencia, también se podría tratar de un virus con ADN cuyo genoma codifica para una enzima que posee de forma accesorial una actividad de transcriptasa inversa como es el caso, por ejemplo, para el virus de la hepatitis B, HBV (12). Además, el carácter aleatorio de los cebadores degenerados usados para esta técnica de amplificación PCR puede haber permitido, debido a las homologías de secuencias imprevistas o de sitios conservados en el gen de una enzima emparentada, la amplificación de un ácido nucleico que procede de un agente patógeno y/o co-infeccioso procariota o eucariota (protista).

EJEMPLO 2: OBTENCIÓN DE CLONES DENOMINADOS MSRV-1B Y MSRV-2B, QUE DEFINEN UNA FAMILIA MSRV-1 Y MSRV-2, MEDIANTE AMPLIFICACIÓN PCR "ANIDADA" DE LAS REGIONES POL CONSERVADAS DE LOS RETROVIRUS, EN PREPARACIONES DE LINFOCITOS B DE UN NUEVO CASO DE SEP

Se ha usado la misma técnica de PCR modificada según la técnica de Shih (12) para amplificar y secuenciar el material nucleico ARN presente en una fracción de viriones purificada en el pico de actividad de transcriptasa

inversa "de tipo LM7", en un gradiente de sacarosa según la técnica descrita por H. Perron (13), y según los protocolos mencionados en el ejemplo 1, a partir de una línea linfoblastoide espontánea, obtenida mediante auto-inmortalización en cultivo de linfocitos B de un paciente SEP seropositivo para el virus de Epstein-Barr (EBV), después del cultivo de las células linfoides sanguíneas en un medio de cultivo apropiado que contiene una concentración apropiada de ciclosporina A. En la figura 4 se ilustra una representación de la actividad de transcriptasa inversa en las fracciones de sacarosa extraídas en un gradiente de purificación de los viriones producidos por esta línea. De la misma manera, se han tratado los sobrenadantes de cultivo de una línea B obtenida en las mismas condiciones a partir de un control libre de SEP en las mismas condiciones, y la determinación de la actividad de transcriptasa inversa en las fracciones del gradiente de sacarosa resultó negativo en todas partes (ruido de fondo) y se representa en la figura 5. Se han analizado la fracción 3 del gradiente que corresponde a la línea B de SEP y la misma fracción sin actividad de transcriptasa inversa del gradiente de control no-SEP mediante la misma técnica de RT-PCR que anteriormente, derivada de Shih (12), seguida de las mismas etapas de clonación y de secuenciación tales como se describen en el ejemplo 1.

Se puede observar que las secuencias de tipo MSRV-1 y MSRV-2 se encuentran en el único material asociado con un pico de actividad de transcriptasa inversa "de tipo LM7" que procede de la línea linfoblastoide B de SEP. Estas secuencias no se han encontrado con el material de la línea linfoblastoide B de control (no-SEP) en 26 clones recombinantes tomados al azar. Sólo se han encontrado las secuencias contaminantes de tipo Mo-MuLV que proceden de la transcriptasa inversa comercial usada para la etapa de síntesis del ADNc y de las secuencias sin analogía retroviral particular en este control, debido a la amplificación "consenso" de las secuencias de polimerasas homólogas que realiza esta técnica de PCR. Además, la ausencia de diana concentrada que compite para la reacción de amplificación en la muestra de control permite la amplificación de contaminantes diluidos. La diferencia de los resultados es de forma evidente altamente significativa (χ^2 , $p < 0,001$).

EJEMPLO 3: OBTENCIÓN DE UN CLON PSJ17, QUE DEFINE UN RETROVIRUS MSRV-1, MEDIANTE REACCIÓN DE TRANSCRIPTASA INVERSA ENDÓGENA EN UNA PREPARACIÓN DE VIRION QUE PROCEDE DE LA LÍNEA PLI-2

Este enfoque prevé obtener secuencias de ADN retro-transcrita a partir del ARN supuesto retroviral en el aislado, usando la actividad de transcriptasa inversa presente en este mismo aislado. Esta actividad de transcriptasa inversa únicamente puede funcionar teóricamente en presencia de un ARN retroviral, enlazado a un cebador ARNt o hibridado con hebras cortas de ADN ya retro-transcritas en las partículas retrovirales (16). Así, se ha optimizado la obtención de secuencias retrovirales específicas en un material contaminado por ácidos nucleicos celulares, según estos autores gracias a la amplificación enzimática específica de las porciones de ARN víricos mediante una actividad de transcriptasa inversa vírica. Para ello, los autores han determinado las condiciones físico-químicas particulares en las que esta actividad enzimática de transcripción inversa en ARN contenidos en viriones podía ser efectiva *in vitro*. Estas condiciones corresponden a la descripción técnica de los protocolos presentados a continuación (reacción de RT endógeno, purificación, clonación y secuenciación).

El enfoque molecular consistió en usar una preparación de virión concentrado, pero no purificado, obtenido a partir de los sobrenadantes de cultivo de la línea PLI-2 preparados según el siguiente método: los sobrenadantes de cultivo se recogen dos veces por semana, se pre-centrifugan a 10.000 r.p.m. durante 30 minutos para eliminar los residuos celulares, y después se congelan a -80°C o se usan tal cual para las siguientes etapas. Los sobrenadantes frescos o descongelados se centrifugan en un tampón de PBS-glicerol al 30% a 100.000 g (o 30.000 r.p.m. en un rotor LKB-HITACHI de tipo 45 T) durante 2 horas a 4°C . Después de la eliminación del sobrenadante, el residuo sedimentado se recoge en un pequeño volumen de PBS y constituye la fracción de virión concentrado, pero no purificado. Esta muestra vírica concentrada pero no purificada se usó para efectuar una reacción denominada de transcripción inversa endógena, tal como se describe a continuación.

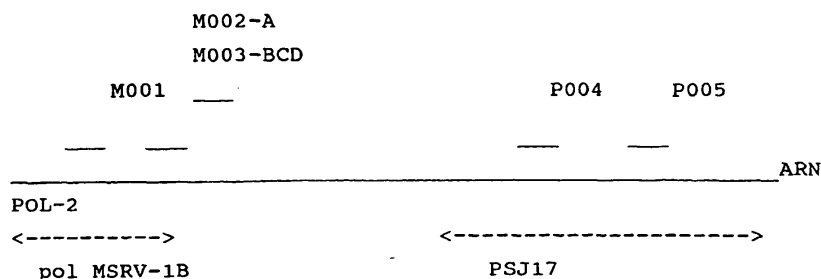
Se descongela un volumen de 200 μl de virión purificado según el protocolo descrito anteriormente y que contiene una actividad de transcriptasa inversa de aproximadamente 1-5 millones de dpm a 37°C hasta aparición de una fase líquida, y después se dispone sobre hielo. Se ha preparado un tampón 5 veces concentrado con los siguientes componentes: se han mezclado 500 mM de Tris-HCl pH 8,2; 75 mM de NaCl; 25 mM de MgCl_2 ; 75 mM de DTT y NP 40 al 0,10%; 100 μl de tampón 5X + 25 μl de una disolución de 100 mM de dATP + 25 μl de una disolución de 100 mM de dTTP + 25 μl de una disolución de 100 mM de dGTP + 25 μl de una disolución de 100 mM de dCTP + 100 μl de agua destilada estéril + 200 μl de la suspensión de viriones (actividad T.I. de 5 millones de DPM) en el PBS y se han incubado a 42°C durante 3 horas. Después de esta incubación, la mezcla de reacción se añade directamente a una mezcla tamponada de fenol/cloroformo/alcohol isoamílico (Sigma ref. P 3803); se recoge la fase acuosa y se añade un volumen de agua destilada estéril a la fase orgánica para volver a extraer el material nucleico residual. Las fases acuosas recogidas se agrupan, y los ácidos nucleicos se precipitan mediante adición de acetato de sodio 3M, pH 5,2 al 1/10 de volumen + 2 volúmenes de etanol + 1 μl de glicógeno (Boehringer-Mannheim ref. 901393) y se ponen a -20°C de la muestra durante 4 horas o una noche a $+4^{\circ}\text{C}$. Después, el precipitado obtenido después de la centrifugación se lava con etanol al 70% y se resuspende en 60 ml de agua destilada. Después, los productos de esta reacción se purificaron, se clonaron y se secuenciaron según el protocolo descrito a continuación: se han generado ADN con extremos libres con adeninas no emparentadas a los extremos: en primer lugar se ha efectuado una reacción de "llenado": se han mezclado 25 μl de la disolución de ADN previamente purificado con 2 μl de una

disolución de 2,5 mM que contiene, en cantidad equimolar, dATP + dGTP + dTTP + dCTP/1 µl de ADN polimerasa T4 (Boehringer-Mannheim ref. 1004 786)/5 µl de 10X "incubation buffer for restriction enzyme" (Boehringer-Mannheim ref. 1417 975)/1 µl de una disolución al 1% de sero-albumina bovina/16 µl de agua destilada estéril. Esta mezcla se incubó durante 20 minutos a 11°C. Se añadieron 50 µl de tampón TE y 1 µl de glicógeno (Boehringer-Mannheim ref. 901 393), antes de la extracción de los ácidos nucleicos con fenol/cloroformo/alcohol isoamílico (Sigma ref. P 3803), y de la precipitación con acetato de sodio como se ha descrito anteriormente. El ADN precipitado después de la centrifugación se resuspende en 10 µl de tampón 10 mM de Tris pH 7,5. Después, se mezclaron 5 µl de esta suspensión con 20 µl de tampón Taq 5X, 20 µl de 5 mM de dATP, 1 µl (5U) de Taq ADN polimerasa (Amplitaq™) y 54 µl de agua destilada estéril. Esta mezcla se incubó durante 2 horas a 75°C con una película de aceite en la superficie de la disolución. El ADN en suspensión en la disolución acuosa extraído debajo de la película de aceite después de la incubación se precipita como se ha descrito anteriormente, y se resuspende en 2 µl de agua destilada estéril. El ADN obtenido se insertó en un plásmido con la ayuda del kit TA Cloning™. Se mezclaron los 2 µl de disolución de ADN 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 "TA DNA LIGASE". Esta mezcla se incubó durante una noche a 12°C. Las siguientes etapas se realizaron según las instrucciones del kit TA Cloning™ (British Biotechnology). Al final del procedimiento, las colonias blancas de bacterias recombinantes (blanco) se replantaron para ser cultivadas y la extracción de los plásmidos incorporados según el procedimiento denominado de "miniprep" (17). La preparación de plásmido de cada colonia recombinante se escindió mediante una enzima de restricción apropiada y se analizó sobre gel de agarosa. Los plásmidos que poseen un inserto detectado bajo luz UV después del marcado del gel con bromuro de etidio se seleccionaron para la secuenciación del inserto, después de la hibridación con un cebador complementario del promotor Sp6 presente en el plásmido de clonación del TA Cloning kit®. Después, se ha realizado la reacción previa a la secuenciación según el método preconizado para el uso del kit de secuenciación "Prism ready reaction kit dye deoxyterminator cycle sequencing kit" (Applied Biosystems, ref. 401384), y se ha realizado la secuenciación automática con el aparato "Automatic Sequencer, modelo 373 A", Applied Biosystems, según las instrucciones del fabricante.

El análisis discriminante en los bancos de datos informáticos de las secuencias clonadas a partir de los fragmentos de ADN presentes en la mezcla de reacción permitió demostrar una secuencia de tipo retroviral. El clon que corresponde a PSJ17 ha sido totalmente secuenciado, y se ha analizado la secuencia obtenida, representada en la figura 6 e identificada mediante la SEC ID nº 9, con la ayuda del programa "Geneworks®" en los bancos de datos actualizados "Genebank®". El análisis de los bancos de datos no ha permitido encontrar ninguna secuencia idéntica ya descrita. Sólo se ha podido encontrar una homología parcial con ciertos elementos retrovirales conocidos. La homología relativa más interesante se refiere a un retrovirus endógeno denominado ERV-9 o HSERV-9, según las referencias (18).

EJEMPLO 4: AMPLIFICACIÓN PCR DE LA SECUENCIA NUCLEICA CONTENIDA ENTRE LA REGIÓN 5' DEFINIDA POR EL CLON "POL MSRV-1B" Y LA REGIÓN 3' DEFINIDA POR EL CLON PSJ17

Se han definido cinco oligonucleótidos, M001, M002, M002-A, M003-BCD, P004 y P005, para amplificar ARN que procede de viriones purificados POL-2. Se han efectuado reacciones de control a fin de controlar la presencia de contaminantes (reacción sobre agua). La amplificación consiste en una etapa de RT-PCR según el protocolo descrito en el Ejemplo 2, seguida de una PCR "anidada" según el protocolo PCR descrito en el documento EP-A-0 569 272. En el primer ciclo de RT-PCR, se usan los cebadores M001 y P004 o P005. En el segundo ciclo de PCR, se usan los cebadores M002-A o M003-BCD, y el cebador P004. Los cebadores se posicionan como sigue:



Su composición es:

cebador M001: GGTCITICCAIGG (SEC ID nº 20)
 cebador M002-A: TTAGGGATAGCCCTCATCTCT (SEC ID nº 21)
 cebador M003-BCD: TCAGGGATAGCCCCATCTAT (SEC ID nº 22)
 cebador P004: AACCTTTGCCACTACATCAATTT (SEC ID nº 23)
 cebador P005: GCGTAAGGACTCCTAGAGCTATT (SEC ID nº 24)

El producto de amplificación "anidada" obtenido y denominado M003-P004 se representa en la figura 7, y

corresponde a la secuencia SEC ID nº 8.

EJEMPLO 5: AMPLIFICACIÓN Y CLONACIÓN DE UNA PORCIÓN DEL GENOMA RETROVÍRICO MSRV-1 CON LA AYUDA DE UNA SECUENCIA YA IDENTIFICADA, EN UNA MUESTRA DE VIRUS PURIFICADA CON PICO DE ACTIVIDAD DE TRANSCRIPTASA INVERSA

Se ha usado una técnica de PCR derivada de la técnica publicada por Frohman (19). La técnica derivada permite, con la ayuda de un cebador específico en dirección 3' del genoma a amplificar, alargar la secuencia hacia la región 5' del genoma a analizar. Esta técnica variante se describe en la bibliografía de la compañía Clontech Laboratories Inc., (Palo-Alto California, USA) suministrada con su producto "5'-AmpliFINDER™ RACE Kit", que se usó en una fracción de virión purificada como se ha descrito anteriormente.

Los cebadores 3' específicos usados en el protocolo del kit para la síntesis del ADNc y la amplificación PCR son respectivamente complementarios a las secuencias MSRV-1 siguientes:

ADNc: TCATCCATGTACCGAAGG(SEC ID nº 25)
amplificación: ATGGGGTTCCAAGTCCCT(SEC ID nº 26)

Los productos que proceden de la PCR se purificaron sobre gel de agarosa según los métodos habituales (17), y después se resuspendieron en 10 ml de agua destilada. Consistiendo una de las propiedades de la polimerasa Taq en añadir una adenina en el extremo 3' de cada una de las dos hebras de ADN, se insertó directamente el ADN obtenido en un plásmido con la ayuda del kit TA Cloning™ (British Biotechnology). Los 2 µl de disolución de ADN se mezclaron con 5 µl de agua destilada estéril, 1 µl de tampón de ligación concentrado 10 veces "10X LIGATION BUFFER", 2 µl de "PCRTM VECTOR" (25 ng/ml) y 1 µl de "TA DNA LIGASE". Esta mezcla se incubó durante una noche a 12°C. Las siguientes etapas se realizaron según las instrucciones del kit TA Cloning® (British Biotechnology). Al final del procedimiento, las colonias blancas de bacterias recombinantes (blanco) se replantaron para ser cultivadas y la extracción de los plásmidos incorporados según el procedimiento denominado "miniprep" (17). La preparación de plásmido de cada colonia recombinante se escindió mediante una enzima de restricción apropiada, y se analizó sobre gel de agarosa. Los plásmidos que poseen un inserto detectado bajo luz UV después del marcado del gel con bromuro de etidio, se seleccionaron para la secuenciación del inserto, después de la hibridación con un cebador complementario del promotor Sp6 presente en el plásmido de clonación del TA Cloning Kit®. Después, se ha realizado la reacción previa a la secuenciación según el método preconizado para el uso del kit de secuenciación "Prism ready reaction kit dye deoxyterminator cycle sequencing kit" (Applied Biosystems, ref. 401384), y se ha realizado la secuenciación automática con el aparato "Automatic sequencer, modelo 373 A" Applied Biosystems según las instrucciones del fabricante.

En primer lugar, esta técnica se ha aplicado a dos fracciones de virión purificado como se describe a continuación, sobre sacarosa a partir del aislado "POL-2" producido por la línea PLI-2 por una parte, y a partir del aislado MS7PG producido por la línea LM7PC por otra parte. Los sobrenadantes de cultivos se recogen dos veces por semana, se precentrifugan a 10.000 r.p.m. durante 30 minutos para eliminar los residuos celulares, y después se congelan hasta -80°C o se usan tal cual para las siguientes etapas. Los sobrenadantes frescos o descongelados se centrifugan sobre un tampón de PBS-glicerol al 30% a 100.000 g (o 30.000 r.p.m. En un rotor LKB-HITACHI de tipo 4,5 T) durante 2 horas a 4°C. Después de la eliminación del sobrenadante, el residuo sedimentado se recoge en un pequeño volumen de PBS y constituye la fracción de viriones concentrados, pero no purificados. Después, el virus concentrado se dispone en un gradiente de sacarosa en un tampón PBS estéril (15 a 50% en peso/peso) y se ultracentrifuga a 35.000 r.p.m. (100.000 g) durante 12 horas a +4°C en un rotor con pocillos. Se recogen 10 fracciones y se extraen 20 µl en cada fracción después de la homogeneización para dosificar la actividad de transcriptasa inversa según la técnica descrita por H. Perron (3). Después, las fracciones que contienen el pico de actividad RT "de tipo LM7" se diluyen en un tampón PBS estéril y se ultracentrifugan durante una hora a 35.000 r.p.m. (100.000 g) para sedimentar las partículas víricas. El residuo de virión purificado así obtenido se recoge entonces en un pequeño volumen de un tampón adecuado para la extracción de ARN. La reacción de síntesis de ADNc mencionada anteriormente se realiza sobre este ARN extraído de virión extracelular purificado. La amplificación de PCR según la técnica citada anteriormente ha permitido obtener el clon F1-11 cuya secuencia identificada por la SEC ID nº 2 se representa en la figura 8.

Este clon permite definir, con los diferentes clones previamente secuenciados, una región de longitud importante (1,2 kb) representativa del gen "pol" del retrovirus MSRV-1, tal como se representa en la figura 9. Esta secuencia denominada SEC ID nº 1 está reconstituida a partir de diferentes clones que se recubren en sus extremos, corrigiendo los artefactos relacionados con los cebadores y con las técnicas de amplificación o de clonación, que interrumpirían artificialmente la trama de lectura del conjunto. Esta secuencia se identificará a continuación con la denominación "región MSRV-1 pol*". Su grado de homología con la secuencia HSERV-9 se representa en la figura 12.

En la figura 9, la trama de lectura potencial con su traducción en aminoácidos se representa debajo de la secuencia de ácidos nucleicos.

EJEMPLO 6: DETECCIÓN DE SECUENCIAS ESPECÍFICAS MSRV-1 Y MSRV-2 EN DIFERENTES MUESTRAS DE PLASMA QUE PROCEDEN DE PACIENTES QUE PADECEN SEP O DE PACIENTES DE CONTROL

Se ha usado una técnica de PCR para detectar los genomas MSERV-1 y MSERV-2 en plasmas obtenidos después de una extracción de sangre sobre EDTA de pacientes que padecen SEP y de pacientes de control no-SEP.

La extracción de los ARN de plasma se ha efectuado según la técnica descrita por P. Chomzynski (20), después de la adición de un volumen del tampón que contiene tiocianato de guanidinio con 1 ml de plasma guardado congelado a -80°C después de la recogida.

Para MSRV-2, la PCR se ha efectuado en las mismas condiciones y con los cebadores siguientes:

- cebador 5', identificado mediante la SEC ID nº 14
5' GTAGTTCGATGTAGAAAGCG 3';

- cebador 3', identificado mediante la SEC ID nº 15
5' GCATCCGGCAACTGCACG 3'.

Sin embargo, también se han obtenido resultados similares con los cebadores de PCR siguientes en dos amplificaciones sucesivas mediante PCR "anidada" sobre muestras de ácidos nucleicos no-tratadas mediante ADNasa.

Los cebadores usados para esta primera etapa de 40 ciclos con una temperatura de hibridación de 48°C son los siguientes:

- cebador 5', identificado mediante la SEC ID nº 27

5' GCCGATATCACCCGCCATGG 3', que corresponde a un cebador PCR MSRV-2 5', para una primera PCR sobre extracción de pacientes,

- cebador 3', identificado mediante la SEC ID nº 28

5' GCATCCGGCAACTGCACG 3', que corresponde a un cebador PCR MSRV-2 3', para una primera PCR sobre extracción de pacientes.

Después de esta etapa, se extraen 10 µl del producto de amplificación y se usan para realizar una segunda amplificación de PCR denominada "anidada" con cebadores situados en el interior de la región ya amplificada. Esta segunda etapa se desarrolla en 35 ciclos, con una temperatura de hibridación de los cebadores ("annealing") de 50°C. El volumen de reacción es de 100 µl.

Los cebadores usados para esta segunda etapa son los siguientes:

- cebador 5', identificado mediante la SEC ID nº 29

5' CGCGATGCTGGTTGGAGAGC 3', que corresponde a un cebador PCR MSRV-2 5', para una PCR-anidada sobre extracción de pacientes,

- cebador 3', identificado mediante la SEC ID nº 30

5' TCTCCACTCCGAATATTCCG 3', que corresponde a un cebador PCR MSRV-2 3', para una PCR-anidada sobre extracción de pacientes.

Para MSRV-1, la amplificación se ha efectuado en dos etapas. Además, la muestra de ácidos nucleicos se trata previamente mediante ADNasa y se efectúa un control de PCR sin RT (transcriptasa inversa AMV) en las dos etapas de amplificación, a fin de verificar que la amplificación de RT-PCR procede exclusivamente del ARN MSRV-1. En el caso de un control sin RT positivo, la muestra alícuota de ARN de partida se trata nuevamente mediante ADNasa, y se amplifica nuevamente.

El protocolo de tratamiento mediante ADNasa desprovista de actividad ARNasa es el siguiente: se alícuota un ARN extraído en presencia de "RNase inhibitor" (Boehringer-Mannheim) en agua tratada con DEPC a una concentración final de 1 µl en 10 µl; a estos 10 µl se añade 1 µl de "RNase-free DNase" (Boehringer-Mannheim) y 1,2 µl de tampón a pH 5 que contiene 0,1 M/l de acetato de sodio y 5 mM/l de MgSO₄, la mezcla se incuba durante 15 minutos a 20°C y se lleva a 95°C durante 1,5 minutos en un "thermocycler".

La primera etapa de RT-PCR MSRV-1 se efectúa según una variante del procedimiento de amplificación de ARN tal como se describe en la solicitud de patente EP-A-0 569 272. En particular, la etapa de síntesis de ADNc se efectúa a 42°C durante una hora; la amplificación de PCR se desarrolla sobre 40 ciclos, con una temperatura de hibridación de los cebadores ("annealing") de 53°C. El volumen de reacción es de 100 µl.

Los cebadores usados para esta primera etapa son los siguientes:

- cebador 5', identificado mediante la SEC ID nº 16
5' AGGAGTAAGGAAACCCAACGGAC 3';

- 5 - cebador 3', identificado mediante la SEC ID nº 17
5' TAAGAGTTGCACAAGTGCG 3'.

Después de esta etapa, se extraen 10 µl del producto de amplificación y se usan para realizar una segunda amplificación de PCR denominada "anidada" con cebadores situados en el interior de la región ya amplificada. Esta segunda etapa se desarrolla sobre 35 ciclos, con una temperatura de hibridación de los cebadores ("annealing") de 53°C. El volumen de reacción es de 100 µl.

Los cebadores usados para esta segunda etapa son los siguientes:

- 15 - cebador 5', identificado mediante la SEC ID nº 18
5' TCAGGGATAGCCCCCATCTAT 3';

- 20 - cebador 3', identificado mediante la SEC ID nº 19
5' AACCCCTTGCCACTACATCAATTT 3'.

En las figuras 10 y 11, se representan los resultados de PCR en forma de fotografía bajo luz ultravioleta de geles de agarosa impregnados de bromuro de etidio, en los que se ha efectuado una electroforesis de los productos de amplificación de PCR depositados separadamente en los diferentes pocillos.

- 25 La fotografía de arriba (figura 10) representa el resultado de la amplificación específica MSRV-1.

El pocillo número 8 contiene una mezcla de marcadores de peso molecular ADN, y los pocillos 1 a 7 representan, en orden, los productos amplificados a partir de los ARN totales de plasmas que proceden de 4 controles sanos libres de SEP (pocillos 1 a 4) y de 3 pacientes que padecen SEP, en diferentes etapas de la enfermedad (pocillos 5 a 7).

En esta serie, se ha detectado material nucleico MSRV-2 en el plasma de un caso de SEP de los 3 ensayados y en ninguno de los 4 plasmas de control. Otros resultados obtenidos en series más extendidas confirman estos resultados.

- 35 La fotografía de abajo (figura 11) representa el resultado de la amplificación específica mediante RT-PCR "anidada" MSRV-1:

el pocillo nº 1 contiene el producto de PCR efectuada con agua sola, sin adición de transcriptasa inversa AMV; el pocillo nº 2 contiene el producto de PCR efectuada con agua sola, sin adición de transcriptasa inversa AMV; el pocillo número 3 contiene una mezcla de marcadores de peso molecular ADN; los pocillos 4 a 13 contienen, en orden, los productos amplificados a partir de los ARN totales extraídos de fracciones de gradiente de sacarosa (recogido de arriba abajo) en el que un residuo de virión que procede de un sobrenadante de cultivo infectado por MSRV-1 y MSRV-2 se ha centrifugado en equilibrio según el protocolo descrito por H. Perron (13); en el pocillo 14 no se ha depositado nada; en los pocillos 15 a 17, se han depositado los productos amplificados de ARN extraído de plasmas que proceden de 3 pacientes diferentes que padecen SEP, en diferentes etapas de la enfermedad.

El genoma retroviral MSRV-1 se encuentra efectivamente en la fracción del gradiente de sacarosa que contiene el pico de actividad de transcriptasa inversa medida según la técnica descrita por H. Perron (3), con una intensidad muy alta (fracción 5 del gradiente, depositada en el pocillo nº 8). Una ligera amplificación ha tenido lugar en la primera fracción (pocillo nº 4) que corresponde posiblemente a un ARN liberado por partículas lisadas que flotaban en la superficie del gradiente; de la misma manera, unos restos agregados han sedimentado en la última fracción (fondo de tubo) arrastrando algunas copias del genoma MSRV-1 que han dado lugar a una amplificación de baja densidad.

En los 3 plasmas de SEP ensayados en esta serie, se ha encontrado el ARN MSRV-1 en un caso, produciendo una amplificación muy intensa (pocillo nº 17).

En esta serie, el genoma ARN retroviral MSRV-1 que corresponde aparentemente a partículas de virus extracelular presentes en el plasma en número extremadamente bajo, se ha detectado mediante RT-PCR "anidada" en un caso de SEP de los 3 ensayados. Otros resultados obtenidos en series más extendidas confirman estos resultados.

Además, la especificidad de las secuencias amplificadas mediante estas técnicas de PCR se pueden verificar y evaluar mediante la técnica "ELOSA" tal como se describe por F. Mallet (21) y en el documento FR-A-2 663 040.

Para MSRV-1, los productos de la PCR anidada descrita anteriormente se pueden ensayar en dos sistemas ELOSA que permiten detectar separadamente un consenso A y un consenso B+C+D de MSRV-1, que corresponden a las

subfamilias descritas en el ejemplo 1 y en las figuras 1 y 2. En efecto, las secuencias cercanas del consenso B+C+D se encuentran esencialmente en las muestras de ARN que proceden de viriones MSRV-1 purificados a partir de cultivos o amplificados en líquidos biológicos extra-celulares de pacientes SEP, mientras que las secuencias cercanas del consenso A se encuentran esencialmente en el ADN celular humano normal.

El sistema ELOSA/MSRV-1 para la captura y la hibridación específica de los productos de PCR de la subfamilia A usa un oligonucleótido de captura cpV1A con un enlace amina en dirección 5', y un oligonucleótido de detección biotinilado dpV1A que tienen respectivamente por secuencia:

- cpV1A identificado mediante la SEC ID nº 31

5' GATCTAGGCCACTTCTCAGGTCCAGS 3', que corresponde al oligonucleótido de captura ELOSA de los productos de PCR-anidada MSRV-1 efectuada con los cebadores identificados mediante las SEC ID nº 16 y SEC ID nº 17, seguida eventualmente de la amplificación con los cebadores identificados mediante las SEC ID nº 18 y SEC ID nº 19 sobre extracción de pacientes;

- dpV1A identificado mediante la SEC ID nº 32

5' CATCTITTTGGICAGGCAITAGC 3' que corresponde al oligonucleótido de captura ELOSA de la subfamilia A de los productos de PCR-"anidada" MSRV-1 efectuada con los cebadores identificados mediante las SEC ID nº 16 y SEC ID nº 17, seguida eventualmente de la amplificación con los cebadores identificados mediante las SEC ID nº 18 y SEC ID nº 19 sobre extracción de paciente.

El sistema ELOSA/MSRV-1 para la captura y la hibridación específica de los productos de PCR de la subfamilia B+C+D usa el mismo oligonucleótido de detección biotinilado dpV1A, y un oligonucleótido de captura cpV1B con un enlace amina en dirección 5' que tienen por secuencia:

- dpV1B identificado mediante la SEC ID nº 33

5' CTTGAGCCAGTTCTCATACCTGGA 3', que corresponde al oligonucleótido de captura ELOSA de la subfamilia B + C + D de los productos de PCR-"anidada" MSRV-1 efectuada con los cebadores identificados mediante las SEC ID nº 16 y SEC ID nº 17, seguida eventualmente de la amplificación con los cebadores identificados mediante las SEC ID nº 18 y SEC ID nº 19 sobre extracción de pacientes.

Este sistema de detección ELOSA ha permitido verificar que ninguno de los productos de PCR así amplificados a partir de plasmas de pacientes SEP tratados mediante ADNasa contenía secuencias de la subfamilia A, y que todos eran positivos con el consenso de las subfamilias B, C y D.

Para MSRV-2, se ha evaluado una técnica ELOSA similar sobre aislados que proceden de cultivos celulares infectados usando los cebadores de amplificación de PCR siguientes:

- cebador 5', identificado mediante la SEC ID nº 34

5' AGTGYTRCCMCARGGCGCTGAA 3', que corresponde a un cebador PCR MSRV-2 5', para PCR sobre extracción de cultivos,

- cebador 3', identificado mediante la SEC ID nº 35

5' GMGGCCAGCAGSAKGTCATCCA 3', que corresponde a un cebador PCR MSRV-2 3', para PCR sobre extracción de cultivos,

y los oligonucleótidos de captura con un enlace amina en dirección 5' cpV2, y de detección biotinilado dpV2, que tienen por secuencias respectivas:

- cpV2 identificado mediante la SEC ID nº 36

5 GGATGCCGCCTATAGCCTCTAC 3', que corresponde a un oligonucleótido de captura ELOSA de los productos de la PCR MSRV-2 efectuada con los cebadores SEC ID nº 34 y SEC ID nº 35, o eventualmente con los cebadores degenerados definidos por Shih (12),

- dpV2 identificado mediante la SEC ID nº 37

5' AAGCCTATCGCGTGCACTTGCC 3', que corresponde a un oligonucleótido de detección ELOSA de los productos de la PCR MSRV-2 efectuada con los cebadores SEC ID nº 34 y SEC ID nº 35, o eventualmente con los cebadores degenerados definidos por Shih (12).

Este sistema de amplificación PCR con un par de cebadores diferente de los que se han descrito anteriormente para la amplificación sobre muestras de pacientes, ha permitido confirmar la infección mediante MSRV-2 de cultivos *in*

vitro y de muestras de ácidos nucleicos usadas para los trabajos de biología molecular.

Por último, los primeros resultados de detección PCR del genoma de agentes patógenos y/o infecciosos, muestran que aparentemente un "virus" libre puede circular en el flujo sanguíneo de pacientes en fase de brote agudo, fuera del sistema nervioso. Esto es compatible con la existencia casi-sistemática de "brechas" en la barrera hemato-encefálica de los pacientes en fase activa de SEP.

EJEMPLO 7: OBTENCIÓN DE SECUENCIAS DEL GEN "env" DEL GENOMA RETROVÍRICO MSRV-1

Tal como ya se ha descrito en el ejemplo 5, se ha usado una técnica de PCR derivada de la técnica publicada por Frohman (19). La técnica derivada permite, con la ayuda de un cebador específico en dirección 3' del genoma a amplificar, alargar la secuencia hacia la región 5' del genoma a analizar. Esta técnica variante se describe en la bibliografía de la compañía "Clontech Laboratories Inc., /Palo-Alto California, USA) suministrada con su producto "5'-AmpliFINDER™ RACE kit", que se ha usado en una fracción de virión purificado como se ha descrito anteriormente.

A fin de realizar una amplificación de la región 3' del genoma retroviral MSRV-1 abarcando la región del gen "env", se ha realizado un estudio para determinar una secuencia consenso en las regiones LTR del mismo tipo que las del retrovirus endógeno defectivo HSERV-9 (18, 24) con el que el retrovirus MSRV-1 presenta homologías parciales.

Se ha usado el mismo cebador 3' específico en el protocolo del kit para la síntesis del ADNc y la amplificación PCR; su secuencia es la siguiente:

GTGCTGATTGGTGTATTTACAATCC (SEC ID nº 45)

Se han realizado en una etapa, la síntesis del ADN complementario (ADNc) y la amplificación PCR monodireccional con el cebador anterior según el procedimiento descrito en la patente EP-A-0 569 272.

Los productos que proceden de la PCR se han extraído después de la purificación sobre gel de agarosa según los métodos convencionales (17), y después se han resuspendido en 10 ml de agua destilada. Consistiendo una de las propiedades de la polimerasa Taq en añadir una adenina en el extremo 3' de cada una de las dos hebras de ADN, el ADN obtenido ha sido directamente insertado en un plásmido con la ayuda del kit TA Cloning™ (British Biotechnology). Los 2 µl de disolución de ADN se han mezclado 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 "TA DNA LIGASE". Esta mezcla se incubó durante una noche a 12°C. Las siguientes etapas se realizaron según las instrucciones del kit TA Cloning® (British Biotechnology). Al final del procedimiento, las colonias blancas de bacterias recombinantes (blanco) se replantaron para ser cultivadas y la extracción de los plásmidos incorporados según el procedimiento denominado "miniprep" (17). La preparación de plásmido de cada colonia recombinante se escindió mediante una enzima de restricción apropiada y se analizó sobre gel de agarosa. Los plásmidos que tienen un inserto detectado bajo luz UV después del marcado del gel con bromuro de etidio, se seleccionaron para la secuenciación del inserto, después de la hibridación con un cebador complementario del promotor Sp6 presente en el plásmido de clonación del TA Cloning kit®. Después, se ha efectuado la reacción previa a la secuenciación según el método preconizado para el uso del kit de secuenciación "Prism ready reaction kit dye deoxyterminator cycle sequencing kit" (Applied Biosystems, ref. 401384), y se ha realizado la secuenciación automática con el aparato "automatic sequencer, modelo 373 A Applied Biosystems según las instrucciones del fabricante.

Este enfoque técnico se ha aplicado a una muestra de virión concentrada como se describe a continuación, a partir de una mezcla de sobrenadantes de cultivo producidos por líneas linfoblastoides B tales como se describen en el ejemplo 2, establecidas a partir de linfocitos de pacientes que padecen SEP y que presentan una actividad de transcriptasa inversa detectable según la técnica descrita por Perron *et al.* (3): los sobrenadantes de cultivos se recogen dos veces por semana, se precentrifugan a 10.000 r.p.m. durante 30 minutos para eliminar los restos celulares, y después se congelan a -80°C o se usan tal cual para las siguientes etapas. Los sobrenadantes frescos o descongelados se centrifugan sobre un tampón de PBS-glicerol al 30% a 100.000 g durante 2 horas a 4°C. Después de la eliminación del sobrenadante, el residuo sedimentado constituye la muestra de viriones concentrados, pero no purificados. El residuo así obtenido se recoge entonces en un pequeño volumen de un tampón adecuado para la extracción de ARN. La reacción de síntesis de ADNc mencionada más arriba se realiza sobre este ARN extraído de virión extracelular concentrado.

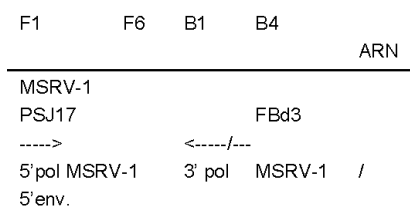
La amplificación de RT-PCR según la técnica citada anteriormente ha permitido obtener el clon FBd3 cuya secuencia identificada mediante la SEC ID nº 46 se representa en la figura 13.

En la figura 14, la homología de secuencia entre el clon FBd3 y el retrovirus HSERV-9 se representa en la tabla matricial mediante un trazo lleno, para cualquier homología parcial superior o igual a 65%. Se puede observar que existen homologías en las regiones flanqueantes del clon (con el gen pol en dirección 5' y con el gen env y después el LTR en dirección 3'), pero que la región interna es totalmente divergente y no presenta ninguna homología, incluso baja, con el gen "env" de HSERV-9. Además, aparece que el clon FBd3 contiene una región "env" más larga que la que se ha descrito para el endógeno defectivo HSERV-9; así, se puede observar que la región divergente

interna constituye un "inserto" entre las regiones de homología parcial con los genes defectivos HSERV-9.

EJEMPLO 8: AMPLIFICACIÓN, CLONACIÓN Y SECUENCIACIÓN DE LA REGIÓN DEL GENOMA RETROVÍRICO MSRV-1 SITUADA ENTRE LOS CLONES PSJ17 Y FBd3

Se han definido cuatro oligonucleótidos, F1, B4, F6 y B1, para amplificar un ARN que procede de viriones concentrados de las cepas POL2 y MS7PG. Se han efectuado reacciones de control de manera que se controla la presencia de contaminantes (reacción sobre agua). La amplificación consiste en una primera etapa de RT-PCR según el protocolo descrito en la solicitud de patente EP-A-0 569 272, seguida de una segunda etapa de PCR efectuada sobre 10 µl de producto de la primera etapa con cebadores internos en la primera región amplificada (PCR "anidada"). En el primer ciclo de RT-PCR, se usan los cebadores F1 y B4. En el segundo ciclo de PCR, se usan los cebadores F6 y B1. Los cebadores están posicionados como sigue:



Su composición es:

cebador F1: TGATGTGAACGGCATACTCACTG (SEC ID nº 47)
 cebador B4: CCCAGAGGTTAGGAAGTCCCTTTC (SEC ID nº 48)
 cebador F6: GCTAAAGGAGACTTGTGGTTGTCAG (SEC ID nº 49)
 cebador B1: CAACATGGGCATTTCCGATTAG (SEC ID nº 50)

El producto de amplificación "anidada" obtenido y denominado "t pol" se representa en la figura 15, y corresponde a la secuencia SEC ID nº 51.

EJEMPLO 9: OBTENCIÓN DE NUEVAS SECUENCIAS, EXPRESADAS EN ARN EN LAS CÉLULAS EN CULTIVO QUE PRODUCEN MSRV-1, Y QUE COMPRENDEN UNA REGIÓN "env" DEL GENOMA RETROVÍRICO MSRV-1

Se ha realizado un banco de ADNc según el procedimiento descrito por el fabricante de los kit "cDNA synthesis module, cDNA rapid adaptator ligation module, cDNA rapid cloning module, lambda gt10 *in vitro* packaging module" (Amersham, ref RPN1256Y/Z, RPN1712, RPN1713, RPN1717, N334Z) a partir del ARN mensajero extraído de células de una línea linfoblástico B tal como se describe en el ejemplo 2, establecida a partir de los linfocitos de un paciente que padece SEP y que presentan una actividad de transcriptasa inversa detectable según la técnica descrita por Perron *et al.* (3).

Se han identificado oligonucleótidos para amplificar un ADNc clonado en el banco nucleico entra la región 3' del clon PSJ17 (pol) y la región 5' (LTR) del clon FBd3. Se han efectuado reacciones de control de manera que se controla la presencia de contaminantes (reacción sobre agua). Las reacciones de PCR efectuadas sobre los ácidos nucleicos clonados en el banco con diferentes pares de cebadores han permitido amplificar una serie de clones que unen unas secuencias pol a las secuencias env o LTR de tipo MSRV-1.

Dos clones son representativos de las secuencias obtenidas en el banco de ADNc celular:

- el clon JLBc1, cuya secuencia SEC ID nº 52 se representa en la figura 16;
- el clon JLBc2, cuya secuencia SEC ID nº 53 se representa en la figura 17.

Las secuencias de los clones JLBc1 y JLBc2 son homólogas a la del clon FBd3, así como aparece en las figuras 18 y 19. La homología entre el clon JLBc1 y el clon JLBc2 se representa en la figura 20.

Las homologías entre los clones JLBc1 y JLBc2 por una parte, y la secuencia HSERV9 por otra parte, se representan respectivamente en las figuras 21 y 22.

Se observa que la región de homología entre JLB1 y JLB2 y FBd3 comprende, con algunas variaciones de secuencia y de tamaño del "inserto", la secuencia suplementaria ausente ("insertada") en la secuencia env HSERV-9, tal como se describe en el ejemplo 8.

Se observa asimismo que la región "pol" clonada es muy homóloga a HSERV-9, que no posee ninguna trama de lectura (teniendo en cuenta los errores de secuencias inducidos por las técnicas usadas, incluyendo el secuenciador

automático) y diverge de las secuencias MSERV-1 obtenidas a partir de viriones. Puesto que estas secuencias se clonaron a partir del ARN de células que expresan partículas MSRV-1, es probable que procedan de elementos retrovíricos endógenos emparentados a la familia ERV9; esto, aun más cuando los genes pol y env están presentes en un mismo ARN que, evidentemente, no es el ARN genómico MSRV-1. Ciertos de estos elementos ERV9 poseen unos LTR funcionales activables mediante virus replicativos que codifican para unos transactivadores homólogos o heterólogos. En estas condiciones, el parentesco entre MSRV-1 y HSERV-9 hace posible la transactivación de los elementos ERV9 endógenos defectivos (o no) por proteínas transactivadores MSRV-1 homólogos, incluso idénticas.

Dicho fenómeno puede inducir una interferencia vírica entre la expresión de MSRV-1 y los elementos endógenos emparentados. Dicha interferencia conduce generalmente a una expresión denominada "defectiva-interferente", de la que ciertas características se han encontrado en los cultivos estudiados infectados por MSRV-1. Además, dicho fenómeno genera la expresión de polipéptidos, incluso de proteínas retrovíricas endógenas que no son obligatoriamente toleradas por el sistema inmunitario. Dicho esquema de expresión aberrante de elementos endógenos emparentados a MSRV-1 e inducido por éste es susceptible de multiplicar los antígenos aberrantes y, por lo tanto, contribuir a la inducción de procesos autoinmunes tales como se observan en la SEP.

Sin embargo, es esencial notar que los clones JLBc1 y JLBc2 difieren de la secuencia ERV9 o HSERV-9 ya descrita, porque poseen una región env más larga que comprende una región suplementaria totalmente divergente de ERV9. Por lo tanto, se puede definir su parentesco con la familia endógena ERV9, pero constituyen evidentemente elementos originales, nunca descritos hasta ahora. En efecto, la interrogación de los bancos de datos de secuencias nucleicas disponibles en la versión nº 15 (1995) del programa "Entrez" (NCBI, NIH, Bethesda, USA) no ha permitido identificar ninguna secuencia homóloga conocida en la región env de estos clones.

EJEMPLO 10: OBTENCIÓN DE LAS SECUENCIAS SITUADAS EN LA REGIÓN 5' pol Y 3' gag DEL GENOMA RETROVÍRICO MSRV-1

Tal como ya se ha descrito en el ejemplo 5, se ha usado una técnica de PCR derivada de la técnica publicada por Frohman (19). La técnica derivada permite, con la ayuda de un cebador específico en dirección 3' del genoma a amplificar, alargar la secuencia hacia la región 5' del genoma a analizar. Esta variante técnica se describe en la documentación de la compañía "Clontech Laboratories Inc., (Palo-Alto California, USA) suministrada con su producto "5'-AmpliFINDER™ RACE Kit", que se ha usado sobre una fracción de virión purificado como se ha descrito anteriormente.

A fin de realizar una amplificación de la región 5' del genoma retrovírico MSRV-1 que empieza en la secuencia pol ya secuenciada (clon F11-1) y que se extiende hacia el gen gag, se han definido unos cebadores específicos MSRV-1.

Los cebadores 3' específicos usados en el protocolo del kit para la síntesis del ADNc y la amplificación PCR son respectivamente complementarios a las secuencias MSRV-1 siguientes:

ADNc: (SEC ID nº 54) CCTGAGTTCTTGCACTAACCC
amplificación: (SEC ID nº 55) GTCCGTTGGGTTTCTTACTCCT

Los productos que proceden de la PCR se han extraído después de la purificación sobre gel de agarosa según los métodos habituales (17), y después se han resuspendido en 10 ml de agua destilada. Consistiendo una de las propiedades de la polimerasas Taq en añadir una adenina en el extremo 3' de cada una de las dos hebras de ADN, el ADN obtenido se ha insertado directamente en un plásmido con la ayuda del kit TA Cloning™ (British Biotechnology). Se han mezclado los 2 µl de disolución de ADN 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 "TA DNA LIGASE". Se ha incubado esta mezcla durante una noche a 12°C. Se han realizado las siguientes etapas según las instrucciones del kit TA Cloning® (British Biotechnology). Al final del procedimiento, las colonias blancas de bacterias recombinantes (blanco) se han recultivado para permitir la extracción de los plásmidos incorporados según el procedimiento denominado "miniprep" (17). La preparación de plásmido de cada colonia recombinante se ha escindido mediante una enzima de restricción apropiada y se ha analizado sobre gel de agarosa. Se han seleccionado los plásmidos que poseen un inserto detectado bajo luz UV después del marcado del gel de bromuro de etidio, para la secuenciación del inserto, después de la hibridación con un cebador complementario del promotor Sp6 presente en el plásmido de clonación del TA Cloning Kit®. Después, se ha efectuado la reacción previa a la secuenciación según el método preconizado para el uso del kit de secuenciación "Prism ready reaction kit dye deoxyterminator cycle Sequencing kit" (Applied Biosystems, ref. 401384) y se ha realizado la secuenciación automática con el aparato "Automatic Sequencer, modelo 373 A Applied Biosystems según las instrucciones del fabricante.

Este enfoque técnico se ha aplicado a una muestra de virión concentrado como se describe a continuación, a partir de una mezcla de sobrenadantes de cultivo producidos mediante líneas linfoblastoides B tales como se describen en el ejemplo 2, establecidas a partir de los linfocitos de pacientes que padecen SEP y que presentan una actividad de transcriptasa inversa detectable según la técnica descrita por Perron *et al.* (3): los sobrenadantes de cultivos se

recogen dos veces por semana, se precentrifugan a 10.000 r.p.m. durante 30 minutos para eliminar los restos celulares, y después se congelan a -80°C o se usan tal cual para las siguientes etapas. Los sobrenadantes frescos o descongelados se centrifugan sobre un tampón de PBS-glicerol 30% a 100.000 g durante 2 horas a 4°C. Después de la eliminación del sobrenadante, el residuo sedimentado constituye la muestra de viriones concentrados, pero no purificados. El residuo así obtenido se recoge entonces en un pequeño volumen de un tampón adecuado para la extracción de ARN. La reacción de síntesis de ADNc mencionada anteriormente se realiza sobre este ARN extraído de virión extracelular concentrado.

La amplificación de RT-PCR según la técnica citada anteriormente, ha permitido obtener el clon GM3 cuya secuencia identificada mediante la SEC ID nº 56 se representa en la figura 23.

En la figura 24, la homología de secuencia entre el clon GMP3 y el retrovirus HSERV-9 se representa en la tabla matricial mediante un trazo lleno, para cualquier homología parcial superior o igual a 65%.

En síntesis, en la figura 25 se representa la localización de los diferentes clones anteriormente estudiados, con relación al genoma ERV9 conocido. En la figura 25, siendo la región env MSRV-1 más larga que el gen env de ERV9 de referencia, la región suplementaria se representa por encima del punto de inserción según una "V", entendiéndose que el material insertado presenta una variabilidad de secuencia y de tamaño entre los clones representados (JLBc1, JLBc2, FBd3). Y en la figura 26 se representa la posición de diferentes clones estudiados en la región MSRV-1 pol*.

Gracias al clon GM3 descrito anteriormente, se ha podido definir una posible trama de lectura, que cubre el conjunto del gen pol, referenciado según la SEC ID nº 57, representado en las figuras sucesivas 27a a 27c.

EJEMPLO 11: DETECCIÓN DE ANTICUERPOS ESPECÍFICOS ANTI-MSRV-1 EN EL SUERO HUMANO

La identificación de la secuencia del gen pol del retrovirus MSRV-1 y de una trama de lectura abierta de este gen ha permitido determinar la secuencia SEC ID nº 39 de aminoácidos de una región de dicho gen, referenciada SEC ID nº 40 (véase la figura 28).

Se han ensayado diferentes péptidos sintéticos que corresponden a fragmentos de la secuencia proteica de la transcriptasa inversa MSRV-1 codificada por el gen pol, para su especificidad antigénica frente a sueros de pacientes que padecen SEP y de pacientes de control sanos.

Se han sintetizado los péptidos químicamente mediante síntesis en fase sólida, según la técnica de Merrifield (Barany G. y Merrifield R.B, 1980, en The Peptides, 2, 1-284, Gross E y Meienhofer J, Eds Academic Press, Nueva York). Las modalidades prácticas son las descritas a continuación.

a) Síntesis de los péptidos:

Se han sintetizado los péptidos sobre una resina de fenilacetamidometilo (PAM)/poliestireno/divinilbenceno (Applied Biosystems, Inc. Foster City, CA) usando un sintetizador automático "Applied Biosystems 430A". Los aminoácidos se enlazan en forma de ésteres de hidroxibenzotriazol (HOBt). Los aminoácidos usados proceden de Novabiochem (Laüflerlingen, Suiza) o de Bachem (Bubendorf, Suiza).

La síntesis química se ha efectuado usando un protocolo de doble enlace con la N-metilpirrolidona (NMP) como disolvente. Los péptidos se han escindido de la resina así como las protecciones laterales, de manera simultánea, con la ayuda de ácido fluorhídrico (HF) en un aparato apropiado (aparato de escisión de tipo I, Peptide Institute, Osaka, Japón).

Para 1 g de peptidilresina, 10 ml de HF, se usan 1 ml de anisol y 1 ml de dimetilsulfuro 5DMS. La mezcla se agita durante 45 minutos a -2°C. Entonces, se evapora el HF a vacío. Después de lavados intensivos con éter, el péptido se eluye de la resina mediante ácido acético al 10%, y después se liofiliza.

Los péptidos se purifican mediante cromatografía líquida ultrarrápida preparativa sobre una columna VYDAC de tipo C18(250 x 21 mm) (The Séparation Group, Hesperia, CA, USA). La elución se realiza mediante un gradiente de acetonitrilo con un caudal de 22 ml/min. Las fracciones recogidas se controlan mediante una elución en condición isocrática sobre una columna VYDAC® C18 analítica (250 x 4,6 mm) con un caudal de 1 ml/min. Las fracciones que presentan el mismo tiempo de retención se reúnen y se liofilizan. Después, se analiza la fracción mayoritaria mediante cromatografía líquida ultrarrápida analítica, con el sistema descrito anteriormente. El péptido que se considera como de pureza aceptable se traduce por un pico único que representa 95% mínimo del cromatograma.

Después, se analizan los péptidos purificados con el objeto de controlar su composición de aminoácidos, con la ayuda de un analizador de aminoácidos automático Applied Biosystems 420H. La medición de la masa molecular química (media) de los péptidos se obtiene usando la espectrometría de masas L.S.I.M.S en modo de ión positivo en un instrumento de doble focalización VG. ZAB.ZSEQ unido a un sistema de adquisición DEC-VAX 2000 (VG

analytical Ltd, Manchester, Inglaterra).

La reactividad de los diferentes péptidos se ha ensayado contra sueros de pacientes que padecen SEP y contra sueros de pacientes de control sanos. Esto ha permitido seleccionar un péptido denominado POL2B cuya secuencia se representa en la figura 28 en el identificador SEC ID nº 39 siguiente, codificado por el gen *pol* de MSRV-1 (nucleótidos 181 a 330).

b) Propiedades antigénicas:

Las propiedades antigénicas del péptido POL2B se han demostrado según el protocolo de ELISA descrito a continuación.

El péptido POL2B liofilizado se ha disuelto en agua destilada estéril con una concentración de 1 mg/ml. Esta disolución madre se ha alicuotado y se ha guardado a +4°C para el uso dentro de una quincena de días, o se ha congelado a -20°C, para un uso dentro de los dos meses. Se ha diluido una alícuota en una disolución de PBS (Phosphate Buffer Saline) a fin de obtener una concentración final de péptido de 1 microgramo/ml. Se disponen 100 microlitros de esta dilución en cada pocillo de placas de microtitulación (plástico "high-binding", COSTAR ref. 3590). Las placas se recubren con un adhesivo de tipo "plate-sealer" y se mantienen durante una noche a +4°C, para la fase de adsorción del péptido sobre el plástico. El adhesivo se quita y las placas se lavan tres veces con un volumen de 300 microlitros de una disolución A (PBS 1x, 0,05% de Tween 20®), después se les da la vuelta sobre una tela absorbente. Las placas así escurridas se llenan con 200 microlitros por pocillo de una disolución B (disolución A + 10% de suero de cabra), y después se recubren con un adhesivo y se incuban durante 45 minutos a 1 hora, a 37°C. Después, las placas se lavan tres veces con la disolución A, como se ha descrito anteriormente.

Las muestras de suero a ensayar se diluyen previamente al 1/50 en la disolución B, y 100 microlitros de cada suero diluido a ensayar se depositan en los pocillos de cada placa de microtitulación. Se deposita un control negativo en un pocillo de cada placa, en forma de 100 microlitros de tampón B. Las placas recubiertas con un adhesivo se incuban entonces durante 1 a 3 horas, a 37°C. Después, las placas se lavan tres veces con la disolución A, como se ha descrito anteriormente. Paralelamente, se diluye un anticuerpo de cabra marcado con peroxidasa y dirigido contra los IgG (Sigma Immunochemicals ref. A6029) o los IgM (Cappel ref. 55228) humanos en la disolución B (dilución 1/5000 para el anti-IgG y 1/1000 para el anti-IgM). Entonces, se depositan 100 microlitros de la dilución adecuada del anticuerpo marcado en cada pocillo de las placas de microtitulación, y las placas recubiertas con un adhesivo se incuban durante 1 a 2 horas, a 37°C. Después, se efectúa nuevamente un lavado de las placas como se ha descrito anteriormente. Paralelamente, el sustrato de la peroxidasa se prepara según las indicaciones del kit "Sigma fast OPD kit" (Sigma Immunochemicals, ref. P9187). Se depositan 100 microlitros de disolución-sustrato en cada pocillo, y las placas se colocan al abrigo de la luz durante 20 a 30 minutos, a temperatura ambiente.

Una vez la reacción coloreada se ha estabilizado, las placas se colocan inmediatamente en un lector espectrofotómetro de placas ELISA, y la densidad óptica (D.O.) de cada pocillo se lee con una longitud de onda de 492 nm. Alternativamente, se depositan 30 microlitros de HCL 1N en cada pocillo para detener la reacción, y las placas se leen mediante el espectrofotómetro en las 24 horas.

Las muestras serológicas se disponen en doble o en triple, y se calcula la densidad óptica (D.O.) que corresponde al suero ensayado calculando la media de las D.O. obtenidas para una misma muestra, con la misma dilución.

La D.O. neta de cada suero corresponde a la D.O. media del suero de la que se resta la D.O. media del control negativo (disolución B: PBS, 0,05% de Tween 20®, 10% de suero de cabra).

c) Detección de anticuerpos IgG anti-MSRV-1 mediante ELISA

Se ha usado la técnica descrita anteriormente con el péptido POL2B para buscar la presencia de anticuerpos IgG específicos anti-MSRV-1 en el suero de 29 pacientes, para los cuales se ha hecho un diagnóstico cierto o probable de SEP según los criterios de Poser (23), y de 32 controles sanos (donantes de sangre).

En la figura 29, se representan los resultados de cada suero ensayado con un anticuerpo anti-IgG. Cada barra vertical representa la densidad óptica (D.O. a 492 nm) neta de un suero ensayado. El eje de ordenadas indica la D.O. neta en el pico de las barras verticales. Las 29 primeras barras verticales situadas a la izquierda de la línea punteada vertical, representan los sueros de 29 casos de SEP ensayados, y las 32 barras verticales situadas a la derecha de la línea punteada vertical, representan los sueros de 32 controles sanos (donantes de sangre).

La media de las D.O. netas de los SEP ensayadas es de: 0,62. La gráfica permite visualizar 5 controles cuya D.O. neta emerge por encima de los valores agrupados de la población de control. Estos valores pueden representar la presencia de IgG específicos en pacientes seropositivos no-enfermos. Por lo tanto, se han evaluado dos métodos para determinar el umbral estadístico de positividad del ensayo.

La media de las D.O. netas de los controles, incluyendo los controles con D.O. netas elevadas, es de 0,36. Sin los 5

controles cuyas D.O. netas son superiores o iguales a 0,5, la media de los controles "negativos" es de 0,33. La desviación típica de los controles negativos es de 0,10. Se puede calcular un umbral de positividad teórica según la fórmula:

- 5 valor umbral (media de las D.O. netas de los controles seronegativos) + (2 ó 3 desviaciones típicas de las D.O. netas de los controles seronegativos)

10 En el primer caso, se considera que existen seropositivos no-enfermos, y el valor umbral es igual a: $0,33 + (2 \times 0,10) = 0,53$. Los resultados negativos representan un "ruido de fondo" no específico de la presencia de anticuerpos específicamente dirigidos contra un epítipo del péptido.

15 En el segundo caso, si el conjunto de los controles constituidos por donantes de sangre con una buena salud aparente se toma como base de referencia, sin excluir los sueros *a priori* seropositivos, la desviación típica de los "controles no-SEP" es de 0,116. Entonces, el valor umbral es: $0,36 + (2 \times 0,116) = 0,59$.

20 Según este análisis, el ensayo es específico de la SEP. En este aspecto, se constata que el ensayo es específico de la SEP, puesto que tal como se muestra en la tabla 1, ningún control tiene una D.O. neta por encima de este umbral. Realmente, este resultado refleja el hecho de que las titulaciones de los anticuerpos en los pacientes que padecen SEP son mayoritariamente más elevadas que en los controles sanos que han estado en contacto con MSRV-1.

TABLA nº 1

	SEP	CONTROLES
	0,681	0,3515
	1,0425	0,56
	0,5675	0,3565
	0,63	0,449
	0,588	0,2825
	0,645	0,55
	0,6635	0,52
	0,576	0,2535
	0,7765	0,55
	0,5745	0,51
	0,513	0,426
	0,4325	0,451
	0,7255	0,227
	0,859	0,3905
	0,6435	0,265
	0,5795	0,4295
	0,8655	0,291
	0,671	0,347
	0,596	0,4495
	0,652	0,3725
	0,602	0,181
	0,525	0,2725
	0,53	0,426
	0,565	0,1915
	0,517	0,222
	0,607	0,395
	0,3705	0,34
	0,397	0,307
	0,4395	0,219
		0,491 0,491
		0,2265
		0,2605
MEDIA	0,62	0,33
DESVIACIÓN TÍPICA	0,14	0,10
VALOR UMBRAL		0,53

25 En función del primer modo de cálculo y como se representa en la figura 29 y en la tabla correspondiente 1, 26 de los 29 sueros SEP dan un resultado positivo (D.O. neta superior o igual a 0,50) indicando la presencia de IgG específicamente dirigidos contra el péptido POL2B, por lo tanto contra una parte de la enzima de transcriptasa inversa del retrovirus MSRV-1 codificada por su gen pol y, por consiguiente, contra el retrovirus MSRV-1. Así, aproximadamente 90% de los pacientes SEP ensayados han reaccionado contra un epítipo portado por el péptido

POL2B y presentan IgG circulantes dirigidos contra este último.

Cinco donantes de sangre, aparentemente con una buena salud, de 32 presentan un resultado positivo. Así, aparece que aproximadamente 15% de la población no-enferma puede haber estado en contacto con un epítipo portado por el péptido POL2B en condiciones que han llevado a una inmunización activa que se traduce por la persistencia de IgG séricos específicos. Estas condiciones son compatibles con una inmunización contra la transcriptasa inversa del retrovirus MSRV-1, durante una infección por el (y/o reactivación del) retrovirus MSRV-1. La ausencia de patología neurológica aparente que evoca la SEP en estos controles seropositivos puede indicar que son portadores sanos, que han eliminado un virus infeccioso después de haberse inmunizado o que constituyen una población con riesgo de portadores crónicos. En efecto, los datos epidemiológicos que muestran que un agente patógeno presente en el entorno de las regiones con alta preponderancia de SEP puede ser la causa de esta enfermedad, implican que una fracción de la población libre de SEP ha estado necesariamente en contacto con dicho agente patógeno. Se ha demostrado que el retrovirus MSRV-1 constituye la totalidad o parte de este "agente patógeno" origen de la SEP, y es por lo tanto normal que unos controles tomados en una población sana presenten anticuerpos de tipo IgG contra componentes del retrovirus MSRV-1. Así, la diferencia de seropreponderancia entre los SEP y la población de control es extremadamente significativa: ensayo "chi-2", $p < 0,001$. Estos resultados están por lo tanto a favor de un papel etiopatogénico de MSRV-1 en la SEP.

d) Detección de anticuerpos IgM anti-MSRV-1 mediante ELISA

Se ha usado la técnica de ELISA con el péptido POL2B para buscar la presencia de anticuerpos específicos IgM anti-MSRV-1 en el suero de 36 pacientes, para los cuales se ha hecho un diagnóstico cierto o probable de SEP según los criterios de Poser (23), y de 42 controles sanos (donantes de sangre).

En la figura 30, se representan los resultados de cada suero ensayado con un anticuerpo anti-IgM. Cada barra vertical representa la densidad óptica (D.O. a 492 nm) neta de un suero ensayado. El eje de ordenadas indica la D.O. neta en el pico de las barras verticales. Las 36 primeras barras verticales situadas a la izquierda de la línea vertical que corta el eje de abscisas representan los sueros de 36 casos de SEP ensayados, y las barras verticales situadas a la derecha de la línea discontinua vertical representan los sueros de 42 controles sanos (donantes de sangre). La línea horizontal trazada en el medio de la gráfica representa un umbral teórico que delimita los resultados positivos (cuyo pico de la barra se sitúa por encima), y los resultados negativos (cuyo pico de la barra se sitúa por debajo).

La media de las D.O. netas de las SEP ensayadas es de: 0,19.

La media de las D.O. netas de los controles es de: 0,09.

La desviación típica de los controles negativos es de: 0,05.

Debido a la pequeña diferencia entre la media y la desviación típica de los controles, el umbral de positividad teórico se puede calcular según la fórmula:

$$\text{valor umbral} = (\text{media de las D.O. netas de los controles seronegativos}) + (3 \times \text{desviación típica de las D.O. netas de los controles seronegativos})$$

El valor umbral es, por lo tanto, igual a $0,09 + (3 \times 0,05) = 0,26$; es decir, en la práctica, 0,25.

Los resultados negativos representan un "ruido de fondo" no específico de la presencia de anticuerpos específicamente dirigidos contra un epítipo del péptido.

Según este análisis y como se muestra en la figura 30 y en la tabla 2 correspondiente, el ensayo IgM es específico de la SEP, puesto que ningún control tiene una D.O. neta por encima del umbral. Siete de los 36 sueros SEP producen un resultado IgM positivo; ahora bien, el estudio de los datos clínicos revela que estos sueros positivos se han extraído durante un primer brote de SEP o de un brote agudo en enfermos no tratados. Se sabe que los IgM dirigidos contra agentes patógenos se producen durante las primeras infecciones o durante las reactivaciones tras una fase de latencia de dicho agente patógeno.

La diferencia de seropreponderancia entre las SEP y la población de control es extremadamente significativa: ensayo "chi-2", $p < 0,001$.

Estos resultados están a favor de un papel etiopatogénico de MSRV-1 en la SEP.

La detección de los anticuerpos IgM e IgG contra el péptido POL2B permite evaluar la evolución de una infección mediante MSRV-1 y/o de la reactivación vírica de MSRV-1.

TABLA nº 2

	SEP	CONTROLES
	0,064	0,243
	0,087	0,11
	0,044	0,098
	0,115	0,028
	0,089	0,094
	0,025	0,038
	0,097	0,176
	0,108	0,146
	0,018	0,049
	0,234	0,161
	0,274	0,113
	0,225	0,079
	0,314	0,093
	0,522	0,127
	0,306	0,02
	0,143	0,052
	0,375	0,062
	0,142	0,074
	0,157	0,043
	0,168	0,046
	1,051	0,041
	0,104	0,13
	0,187	0,153
	0,044	0,107
	0,053	0,178
	0,153	0,114
	0,07	0,078
	0,033	0,118
	0,104	0,177
	0,187	0,026
	0,044	0,024
	0,053	0,046
	0,153	0,116
	0,07	0,04
	0,033	0,028
	0,973	0,073
		0,008
		0,074
		0,141
		0,219
		0,047
		0,017
MEDIA	0,19	0,09
DESVIACIÓN TÍPICA	0,23	0,05
VALOR UMBRAL		0,26

e) Búsqueda de epítomos inmunodominantes en el péptido POL2B:

5

Con el fin de reducir el ruido de fondo no específico y de optimizar la detección de las respuestas de los anticuerpos anti-MSRV-1, se ha realizado la síntesis de octapéptidos, sucesivamente desplazados de un aminoácido, que cubren el conjunto de la secuencia determinada mediante POL2B, según el protocolo descrito a continuación.

10

Se ha realizado la síntesis química de octapéptidos solapantes que cubren la secuencia de aminoácidos 61-110 representada en el identificador SEC ID nº 39 en una membrana de celulosa activada según la técnica de BERG *et al.* (1989. J. Ann. Chem. Soc., 111, 8024-8026) comercializada por Cambridge Research Biochemicals con la denominación comercial Spotskan. Esta técnica permite la síntesis simultánea de un gran número de péptidos y su análisis.

15

La síntesis se realiza con aminoácidos esterificados cuyo grupo α -amina está protegido por un grupo FMOC (Nova

Biochem), y los grupos laterales por grupos protectores tales como tritilo, éster de t-butilo o éter de t-butilo. Los aminoácidos se solubilizan en N-metil-pirrolidona (NMP) a la concentración de 300 nM, y se depositan 0,9 µl a nivel de los puntos de depósito de azul de bromofenol. Después de una incubación de 15 minutos, se realiza nuevamente un depósito de aminoácidos seguido de otra incubación de 15 minutos. Si se efectúa el enlace entre dos aminoácidos correctamente, se observa una modificación de la coloración (paso del azul al verde-amarillo). Después de tres lavados en DMF, se efectúa una etapa de acetilación mediante anhídrido acético. Se desprotegen a continuación los grupos aminados terminales de los péptidos en curso de síntesis mediante piperidina al 20% en DMF. Los puntos de depósito se recolorean mediante una disolución de azul de bromofenol al 1% en DMF, se lavan tres veces con metanol y se secan. El conjunto de estas operaciones constituye un ciclo de adición de un aminoácido, y este ciclo se repite hasta el final de la síntesis. Cuando se han añadido todos los aminoácidos, se desprotege el grupo NH₂-terminal del último aminoácido mediante piperidina al 20% en DMF, y se acetila mediante anhídrido acético. Los grupos protectores de la cadena lateral se retiran mediante una mezcla de diclorometano/ácido trifluoroacético/triisobutilsilano (5 ml/5 ml/250 µl). Después se ensaya la inmunorreactividad de los péptidos mediante ELISA.

Después de la doble síntesis de los diferentes octapéptidos sobre dos membranas diferentes, estas últimas se aclaran con metanol y se lavan en TBS (Tris 0,1M pH 7,2), y después se incuban durante una noche a temperatura ambiente en un tampón de saturación. Después de varios lavados en TBS-T (tris 0,1M pH 7,2 - 0,05% de Tween 20), se incuba una membrana con una disolución al 1/50 de un suero de referencia que procede de un paciente que padece SEP, y la otra membrana con una dilución al 1/50 de un grupo de sueros de controles sanos. Las membranas se incuban durante 4 horas a temperatura ambiente. Después de los lavados con TBS-T, se añade un conjugado anti-inmunoglobulinas humanas marcado con β-galactosidasa (comercializado por Cambridge Research Biomedicals) a una dilución al 1/200, y el conjunto se incuba durante dos horas a temperatura ambiente. Después de los lavados de las membranas con TBS-T 0,05% y PBS, se revela la inmunorreactividad a nivel de los diferentes puntos mediante adición de 5-bromo-4-cloro-3-indoil-β-D-galactopiranosida en potasio. La intensidad de coloración de los puntos se estima cualitativamente con un valor relativo de 0 a 5, como se representa en las figuras 31 a 33 anexas.

Así, se puede determinar dos regiones inmunodominantes en cada extremo del péptido POL2B que corresponden respectivamente a las secuencias de aminoácidos 65-75 (SEC ID nº 41) y 92-109 (SEC ID nº 42), según la figura 34, y comprendidas respectivamente entre los octapéptidos Phe-Cys-Ile-Pro-Val-Arg-Pro-Asp (FCIPVRPD) y Arg-Pro-Asp-Ser-Gln-Phe-Leu-Phe (RPDSQFLF), y Thr-Val-Leu-Pro-Gln-Gly-Phe-Arg (TVLPQGFR) y Leu-Phe-Gly-Gln-Ala-Leu-Ala-Gln (LFGQALAQ) y una región menos activa, pero aparentemente más específica, debido a que no produce ningún ruido de fondo con el suero de control, representada mediante los octapéptidos Leu-Phe-Ala-Phe-Glu-Asp-Pro-Leu (LFAFEDPL) (SEC ID nº 43) y Phe-Ala-Phe-Glu-Asp-Pro-Leu-Asn (FAFEDPLN) (SEC ID nº 44).

Estas regiones permiten definir nuevos péptidos más específicos y más inmunorreactivos según las técnicas habituales.

Así, es posible, gracias a los descubrimientos efectuados y a los métodos llevados a cabo por los inventores, realizar un diagnóstico de la infección y/o de la reactivación MSRV-1, y evaluar una terapia en la SEP en base a su eficacia para "hacer negativa" la detección de estos agentes en los fluidos biológicos de los pacientes. Además, la detección precoz en personas que no presentan todavía señales neurológicas de SEP, podría permitir instaurar un tratamiento tanto más eficaz de la evolución clínica posterior por cuanto que precedería el estado lesional que corresponde a la aparición de los trastornos neurológicos. Ahora bien, hoy en día, no se puede establecer un diagnóstico de SEP antes de la instalación de una sintomatología neurológica lesional y, por lo tanto, no se ha instaurado ningún tratamiento antes de la emergencia de una clínica evocadora de lesiones del sistema nervioso central ya consecuentes. El diagnóstico de una infección y/o de una reactivación de MSRV-1 y/o MSRV-2 en el ser humano es por lo tanto determinante, y la presente invención suministra los medios para ello.

Así, es posible, además de realizar un diagnóstico de la infección y/o de la reactivación de MSRV-1, evaluar una terapia en la SEP en base a su eficacia en "hacer negativa" la detección de estos agentes en los fluidos biológicos de los pacientes.

EJEMPLO 12: OBTENCIÓN DE UN CLON LB19 QUE CONTIENE UNA PARTE DEL GEN GAG DEL RETROVIRUS MSRV-1

Se ha usado una técnica de PCR derivada de la técnica publicada por Gonzalez-Quintal R. *et al.* (19) y PLAZA *et al.* (25). A partir de los ARN totales extraídos de una fracción de virión purificado como se ha descrito anteriormente, se ha sintetizado el ADNc con la ayuda de un cebador específico (SEC ID nº 64) en dirección 3' del genoma a amplificar, usando la EXPANDTM REVERSE TRANSCRIPTASE (BOEHRINGER MANNHEIM).

ADNc: AAGGGGCATG GACGAGGTGG TGGCTTATTT (SEC ID nº 65) (anti-sentido)

Después de la purificación, se ha añadido una cola poli G en el extremo 5' del ADNc con la ayuda del kit "Terminal transferases kit" comercializado por la compañía Boehringer Mannheim, según el protocolo del fabricante.

Se ha realizado una PCR de anclaje con la ayuda de los cebadores 5' y 3' siguientes:

5 AGATCTGCAG AATTCGATAT CACCCCCCCC CCCCCC (SEC ID nº 91) (sentido),
AAATGTCTGC GGCACCAATC TCCATGTT (SEC ID nº 64) (anti-sentido)

Después, se ha realizado una PCR de anclaje semi-anidada con los cebadores 5' y 3' siguientes:

10 AGATCTGCAG AATTCGATAT CA (SEC ID nº 92) (sentido), y
AAATGTCTGC GGCACCAATC TCCATGTT (SEC ID nº 64) (anti-sentido)

Los productos que proceden de la PCR se han purificado después de la purificación sobre gel de agarosa según los métodos habituales (17), y después se han resuspendido en 10 microlitros de agua destilada. Consistiendo una de las propiedades de la polimerasa Taq en añadir una adenina en el extremo 3' de cada una de las dos hebras de ADN, el ADN obtenido se ha insertado directamente en un plásmido con la ayuda del kit TA Cloning™ (British Biotechnology). Se han mezclado los 2 µl de disolución de ADN 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 "T4DNA LIGASE". Esta mezcla se ha incubado durante una noche a 12°C. Las siguientes etapas se han realizado según las instrucciones del kit TA Cloning® (British Biotechnology). Al final del procedimiento, las colonias blancas de bacterias recombinantes (blanco) se han replantado para ser cultivadas y permitir la extracción de los plásmidos incorporados según el procedimiento denominado "miniprep" (17). La preparación de plásmido de cada colonia recombinante se ha escindido mediante una enzima de restricción apropiada, y se ha analizado sobre gel de agarosa. Los plásmidos que poseen un inserto detectado bajo luz UV después del marcado del gel con bromuro de etidio, se han seleccionado para la secuenciación del inserto, después de la hibridación con un cebador complementario del promotor Sp6 presente en el plásmido de clonación del TA Cloning kit®. Después se ha efectuado la reacción previa a la secuenciación según el método preconizado para el uso del kit de secuenciación "Prism ready reaction kit dye deoxyterminator cycle sequencing kit" (Applied Biosystems, ref. 401384), y se ha realizado la secuenciación automática con el aparato "Automatic sequencer, modelo 373 A" Applied Biosystems según las instrucciones del fabricante.

30 Se ha usado la amplificación PCR según la técnica citada anteriormente sobre un ADNc sintetizado a partir de los ácidos nucleicos de fracciones de partículas infecciosas purificadas en un gradiente de sacarosa, según la técnica descrita por H. Perron (13), a partir de sobrenadantes de cultivo de linfocitos B de un paciente que padece SEP, inmortalizado por la cepa B95 del virus de Epstein-Barr (EBV), y que expresa unas partículas retrovíricas asociadas con una actividad de transcriptasa inversa tal como la descrita por Perron *et al.* (3) y en las solicitudes de patentes francesas SEP 10, 11 y 12. El clon LB19 cuya secuencia identificada mediante SEC ID nº 59 se representa en la figura 35.

40 Este clon permite definir, con el clon GM3 previamente secuenciado y el clon G+E+A (véase el Ejemplo 15), una región de 690 pares de bases representativa de una parte significativa del gen gag del retrovirus MSRV-1, tal como se representa en la figura 36. Esta secuencia denominada SEC ID nº 88 está reconstituida a partir de diferentes clones que se solapan en sus extremos. Esta secuencia se identifica con la denominación región MSRV-1 "gag*". En la figura 36, se representa una trama de lectura potencial con la traducción en aminoácidos por debajo de la secuencia de ácidos nucleicos.

45 **EJEMPLO 13: OBTENCIÓN DE UN CLON FBd13 QUE CONTIENE UNA REGIÓN DE GENE *pol*/ EMPARENTADA AL RETROVIRUS MSRV-1, Y UNA REGIÓN ENV APARENTEMENTE INCOMPLETA QUE CONTIENE UNA TRAMA DE LECTURA (ORF) POTENCIAL PARA UNA GLUCOPROTEÍNA**

50 Extracción de los ARN víricos: Se han extraído los ARN según el método brevemente descrito a continuación.

Se centrifuga un "grupo" de sobrenadante de cultivo de linfocitos B de pacientes que padecen SEP (650 ml) durante 30 minutos a 10.000 g. El residuo vírico obtenido se resuspende en 300 microlitros de PBS/10 mM de MgCl₂. El material se trata mediante una mezcla ADNasa (100 mg/ml)/ARNasa (50 mg/ml) durante 30 minutos a 37°C, y después mediante proteinasa K (50 mg/ml) durante 30 minutos a 46°C.

Se extraen los ácidos nucleicos mediante un volumen de una mezcla de fenol/SDS al 0,1% (v/v) calentado a 60°C, y después se extraen de nuevo mediante un volumen de fenol/cloroformo (1/1; v/v).

60 La precipitación del material se efectúa mediante 2,5 V de etanol en presencia de 0,1 V de acetato de sodio pH = 5,2. El residuo obtenido después de la centrifugación se resuspende en 50 microlitros de agua DEPC estéril.

La muestra se trata nuevamente mediante 50 mg/ml de ADNasa "ARNase free" durante 30 minutos a temperatura ambiente, se extrae mediante un volumen de fenol/cloroformo, y se precipita en presencia de acetato de sodio y de etanol.

El ARN obtenido se cuantifica mediante lectura de D.O. a 260 nm. La presencia de MSRV-1 y la ausencia de contaminante ADN se controlan mediante una PCR y una RTPCR MSRV-1 específica asociada a una ELOSA específica del genoma MSRV-1.

5 Síntesis de ADNc:

Se usan 5 mg de ARN para sintetizar un ADNc cebado mediante un oligonucleótido polidT según las instrucciones del kit "cDNA Synthesis Module" (ref. RPN 1256, Amersham) con algunas modificaciones: la retrotranscripción se efectúa a 45°C en lugar de los 42°C recomendados.

El producto de síntesis se purifica mediante una doble extracción y una doble purificación siguiendo las instrucciones del fabricante.

La presencia de MSRV-1 se verifica mediante una PCR MSRV-1 asociada con una ELOSA específica del genoma MSRV-1.

"Long distance PCR": (LD-PCR)

Se usan 500 ng de ADNc para la etapa de LD-PCR (Expand Long Template System; Boehringer (ref. 1681842)).

Se han usado varios pares de oligonucleótidos. Entre éstos, el par definido por los cebadores siguientes:

cebador 5': GGAGAAGAGC AGCATAAGTG G(SEC ID nº 66)
cebador 3': GTGCTGATTG GTGTATTTAC AATCC(SEC ID nº 67).

Las condiciones de amplificación son las siguientes:

94°C durante 10 segundos
56°C durante 30 segundos
68°C durante 5 minutos;

Diez ciclos y después 20 ciclos con un incremento de 20 segundos a cada ciclo sobre el tiempo de elongación. Al final de esta primera amplificación, se someten 2 microlitros de producto de amplificación a una segunda amplificación en las mismas condiciones que las anteriores.

Las LD-PCR se llevan a un aparato PCR Perkin modelo 9600 en microtubos de pared fina (Boehringer).

Los productos de amplificación se controlan mediante electroforesis de 1/5 del volumen de amplificación (10 microlitros) en gel de agarosa al 1%. Para el par de cebadores descrito anteriormente, se obtiene una banda de aproximadamente 1,7 Kb.

Clonación del fragmento amplificado:

El producto de PCR se ha purificado mediante el paso sobre un gel de agarosa preparativo y después sobre una columna Costar (Spin; D. Dutcher) según las instrucciones del proveedor.

Se empalman 2 microlitros de la disolución purificada con 50 ng de vector PCRII según las instrucciones del proveedor (TA Cloning kit; British Biotechnology).

El vector recombinante obtenido se aísla mediante transformación de las bacterias DH5aF' competentes. Las bacterias se seleccionan según su resistencia a la ampicilina y la pérdida del metabolismo para Xgal (= colonias blancas). La estructura molecular del vector recombinante se confirma mediante una minipreparación plasmídica, y se hidroliza mediante la enzima *EcoR1*.

Se ha seleccionado FBd13, un clon positivo para todos estos criterios. Se ha efectuado una preparación a gran escala del plásmido recombinante con la ayuda del kit Midiprep Quiagen (ref. 12243) según las instrucciones del proveedor.

La secuenciación del clon FBd3 se efectúa gracias al kit Prism Ready Amplitaq FS dye terminador Perkin (ref. 402119) siguiendo las instrucciones del fabricante. Las reacciones de secuencias se depositan en un secuenciador automático Perkin de tipo 377 ó 373A. La estrategia de secuenciación consiste en un "paseo sobre el gen" realizado sobre las dos hebras del clon FBd13.

La secuencia del clon FBd13 se identifica mediante la SEC ID nº 58.

En la figura 37, la homología de secuencia entre el clon FBd13 y el retrovirus HSERV-9 se representa en la tabla

matricial mediante un trazo lleno, para cualquier homología parcial superior o igual a 70%. Se puede constatar que existen homologías en las regiones flanqueantes del clon (con el gen *pol* en dirección 5' y con el gen *env*, y después LTR en dirección 3'), pero que la región interna es totalmente divergente y no presenta ninguna homología, incluso baja, con el gen *env* de HSERV-9. Además, aparece que el clon FBd13 contiene una región "env" más larga que la que se describe para el endógeno defectivo HSERV-9; así, se puede constatar que la región divergente interna constituye un "inserto" entre las regiones de homología parcial con los genes defectivos HSERV-9.

Esta secuencia suplementaria determina una orf potencial, denominada ORF B13, representada mediante su secuencia de aminoácido SEC ID nº 87.

La estructura molecular del clon FBd13 se ha analizado con la ayuda del programa GeneWork y de los bancos de datos Genbank y SwissProt.

Se han encontrado 5 sitios de glucosilación.

La proteína no tiene ninguna homología significativa con unas secuencias ya conocidas.

Es probable que es clon proceda de una recombinación con un elemento retrovítico endógeno (ERV), relacionado con la replicación de MSRV-1.

Dicho fenómeno genera la expresión de polipéptidos, incluso de proteínas retrovíticas endógenas que no están necesariamente toleradas por el sistema inmunitario. Dicho esquema de expresión aberrante de elementos endógenos emparentados a MSRV-1 y/o inducida por éste es susceptible de multiplicar los antígenos aberrantes y, por lo tanto, contribuir a la inducción de procesos autoinmunes tales como se observan en la SEP. Evidentemente, constituye un elemento original nunca descrito hasta ahora. En efecto, la interrogación de los bancos de datos de secuencias nucleicas disponibles en la versión nº 19 (1996) del programa "Entrez" (NCBI, NIH, Bethesda, USA) no ha permitido identificar ninguna secuencia homóloga conocida que comprende el conjunto de la región env de este clon.

EJEMPLO 14: OBTENCIÓN DE UN CLON FP6 QUE CONTIENE UNA PARTE DEL GEN *pol* CON UNA REGIÓN QUE CODIFICA PARA LA ENZIMA DE TRANSCRIPTASA INVERSA HOMÓLOGA DEL CLON POL*MSRV-1, Y UNA REGIÓN 3'*pol* DIVERGENTE DE LAS SECUENCIAS EQUIVALENTES DESCRITAS EN LOS CLONES POL*, tpol, FBd3, JLBc1 Y JLBc2

Se ha efectuado una 3'RACE sobre un ARN total extraído de plasma de un paciente que padece SEP. Se ha usado como control negativo un plasma de control sano, tratado con las mismas condiciones. Se ha realizado la síntesis de ADNc con el cebador oligo dT modificado siguiente:

5' GACTCGCTGC AGATCGATTT TTTTTTTTTT TTTT 3' (SEC ID nº 68)

y la transcriptasa inversa "Expand RT" de Boehringer según las condiciones preconizadas por la compañía. Se ha efectuado una PCR con la enzima KlenTaq (Clontech) con las siguientes condiciones: 94°C durante 5 minutos y después 93°C durante 1 minuto, 58°C durante 1 minuto, 68°C durante 3 minutos durante 40 ciclos y 68°C durante 8 minutos y con un volumen de reacción final de 50 µl.

Cebadores usados para la PCR:

-cebador 5', identificado mediante la SEC ID nº 69

5' GCCATCAAGC CACCCAAGAA CTCTTAATT 3';

-cebador 3', identificado mediante la SEC ID nº 68 (= la misma que para el ADNc)

Se ha realizado una segunda PCR denominada "semi-anidada" con un cebador 5' situado en el interior de la región ya amplificada. Esta segunda PCR se ha efectuado en las mismas condiciones experimentales que las usadas durante la primera PCR, usando 10 µl del producto de amplificación que procede de la primera PCR.

Cebadores usados para la PCR semi-anidada:

- cebador 5', identificado mediante la SEC ID nº 70

5' CCAATAGCCA GACCATTATA TACTTAATT 3';

- cebador 3', identificado mediante la SEC ID nº 68 (= la misma que para el ADNc)

Los cebadores SEC ID nº 69 y SEC ID nº 70 son específicos de la región pol*: posición nº 403 a nº 422 y nº 641 a nº 670 respectivamente.

Así, se ha obtenido un producto de amplificación a partir del ARN extracelular extraído del plasma de un paciente que padece SEP. El fragmento correspondiente no se ha observado para el plasma de un control sano. Este

producto de amplificación se ha clonado de la siguiente manera.

El ADN amplificado se insertó en un plásmido con la ayuda del kit TA Cloning™. Se mezclaron los 2 µl de disolución de ADN 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 "TA DNA LIGASE". Esta mezcla se incubó durante una noche a 12°C. Las siguientes etapas se realizaron según las instrucciones del kit TA Cloning™ (British Biotechnology). Al final del procedimiento, las colonias blancas de bacterias recombinantes (blanco) se replantaron para ser cultivadas y la extracción de los plásmidos incorporados según el procedimiento denominado de "miniprep" (17). La preparación de plásmido de cada colonia recombinante se escindió mediante una enzima de restricción apropiada y se analizó sobre gel de agarosa. Los plásmidos que poseen un inserto detectado bajo luz UV después del marcado del gel con bromuro de etidio se seleccionaron para la secuenciación del inserto, después de la hibridación con un cebador complementario del promotor Sp6 presente en el plásmido de clonación del TA Cloning kit®. Después, se ha realizado la reacción previa a la secuenciación según el método preconizado para el uso del kit de secuenciación "Prism ready reaction kit dye deoxyterminator cycle sequencing kit" (Applied Biosystems, ref. 401384), y se ha realizado la secuenciación automática con el aparato "Automatic Sequencer, modelo 373 A" Applied Biosystems, según las instrucciones del fabricante.

El clon obtenido, denominado FP6, permite definir una región de 467 pb homóloga en 89% a la región pol* del retrovirus MSRV-1, y una región de 1167 pg homóloga en 64% a la región pol de ERV-9 (nº 1634 a 2856).

El clon FP6 se representa en la figura 38 mediante su secuencia nucleotídica identificada mediante la SEC ID nº 61. Las tres tramas de lectura potenciales de este clon se representan mediante su secuencia de aminoácido debajo de la secuencia nucleotídica.

EJEMPLO 15: OBTENCIÓN DE UNA REGIÓN DENOMINADA G+E+A QUE CONTIENE UNA ORF PARA UNA PROTEASA RETROVÍRICA MEDIANTE AMPLIFICACIÓN PCR DE LA SECUENCIA NUCLEICA CONTENIDA ENTRE LA REGIÓN 5' DEFINIDA POR EL CLON "GM3" Y LA REGIÓN 3' DEFINIDA POR EL CLON POL*, A PARTIR DEL ARN EXTRAÍDO DE UN GRUPO DE PLASMAS DE PACIENTES QUE PADECEN SEP

Se ha definido oligonucleótidos específicos de las secuencias MSRV-1 ya identificados por el solicitante para amplificar el ARN retrovítico de viriones presentes en el plasma de pacientes que padecen SEP. Se han efectuado reacciones de control a fin de controlar la presencia de contaminantes (reacción sobre agua). La amplificación consiste en una etapa de RT-PCR, seguida de una PCR "anidada". Se han definido pares de cebadores para amplificar tres regiones solapantes (denominadas G, E y A) en las regiones definidas por las secuencias de los clones GM3 y pol*, previamente descritos.

RT-PCR semi-anidada para la amplificación de la región G:

- en el primer ciclo RT-PCR, se usan los siguientes cebadores:

cebador 1: SEC ID nº 71 (sentido)
cebador 2: SEC ID nº 72 (anti-sentido)

- en el segundo ciclo PCR, se usan los siguientes cebadores:

cebador 1: SEC ID nº 73 (sentido)
cebador 4: SEC ID nº 74 (anti-sentido)

RT-PCR anidada para la amplificación de la región E:

- en el primer ciclo RT-PCR, se usan los siguientes cebadores:

cebador 5: SEC ID nº 75 (sentido)
cebador 6: SEC ID nº 76 (anti-sentido)

- en el segundo ciclo PCR, se usan los siguientes cebadores:

cebador 7: SEC ID nº 77 (sentido)
cebador 8: SEC ID nº 78 (anti-sentido)

RT-PCR semi-anidada para la amplificación de la región A:

- en el primer ciclo RT-PCR, se usan los siguientes cebadores:

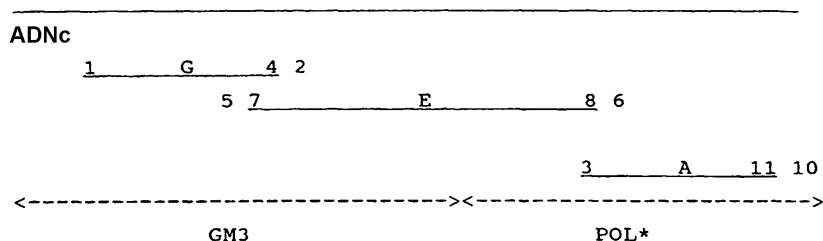
cebador 9: SEC ID nº 79 (sentido)
cebador 10: SEC ID nº 80 (anti-sentido)

- en el segundo ciclo PCR, se usan los siguientes cebadores:

cebador 9: SEC ID nº 81 (sentido)

cebador 11: SEC ID nº 82 (anti-sentido)

Los cebadores y las regiones G, E y A que definen se posicionan como en lo siguiente:



Se ha determinado la secuencia de la región definida por los diferentes clones G, E y A después de la clonación y de la secuenciación de los productos de amplificación "anidada".

Se han reunido los clones G, E y A mediante PCR con los cebadores 1 en dirección 5' del fragmento G y 11 en dirección 3' del fragmento A, descritos anteriormente. Se ha amplificado un fragmento G+E+A de aproximadamente 1580 pb y se ha insertado en un plásmido con la ayuda del kit TA Cloning (marca comercial). Se ha determinado la secuencia del producto de amplificación que corresponde a G+E+A y se ha realizado el análisis de los recubrimientos G+E y E+A. La secuencia se representa en la figura 39 y corresponde a la secuencia SEC ID nº 89.

Se ha encontrado una trama de lectura que codifica para una proteasa retroviral MSR-1 en la región E. La secuencia de aminoácido de la proteasa, identificada mediante SEC ID nº 90, se representa en la figura 40.

EJEMPLO 16: OBTENCIÓN DE UN CLON LTRGAG12, EMPARENTADO A UN ELEMENTO RETROVIRICO ENDÓGENO (ERV) CERCANO DE MSR-1, EN EL ADN DE UNA LÍNEA LINFOBLASTOIDE DE SEP QUE PRODUCE VIRIONES Y QUE EXPRESA EL RETROVIRUS MSR-1

Se ha efectuado una PCR anidada sobre el ADN extraído de una línea linfoblastoide (linfocitos B inmortalizados por el virus EBV, cepa B95, como se ha descrito anteriormente y como es bien conocido por el experto en la materia) que expresa el retrovirus MSR-1 y que procede de los linfocitos de la sangre periférica de un enfermo que padece SEP.

En la primera etapa de PCR, se han usado los siguientes cebadores:

cebador 4327: CTCGATTCTTGCTGGGCCTTA (SEC ID nº 83)

cebador 3512: GTTGATTCCC TCCTCAAGCA (SEC ID nº 84)

Esta etapa comprende 35 ciclos de amplificación con las siguientes condiciones: 1 minuto a 94°C, 1 minuto a 54°C y 4 minutos a 72°C.

En la segunda etapa de PCR, se han usado los siguientes cebadores:

cebador 4294: CTCTACCAAT CAGCATGTGG (SEC ID nº 85)

cebador 3591: TGTTCTCTT GGTCCCTAT (SEC ID nº 86)

Esta etapa comprende 35 ciclos de amplificación con las siguientes condiciones: 1 minuto a 94°C, 1 minuto a 54°C y 4 minutos a 72°C.

Los productos que proceden de la PCR se han purificado después de la purificación sobre gel de agarosa según los métodos convencionales (17), y después se han resuspendido en 10 ml de agua destilada. Consistiendo una de las propiedades de la polimerasa Taq en añadir una adenina en el extremo 3' de cada una de las dos hebras de ADN, el ADN obtenido ha sido directamente insertado en un plásmido con la ayuda del kit TA CloningTM (British Biotechnology). Los 2 µl de disolución de ADN se han mezclado con 5 µl de agua destilada estéril, 1 µl de un tampón de ligación 10 veces concentrado "10X LIGATION BUFFER", 2 µl de "pCRTM VECTOR" (25 ng/ml) y 1 µl de "TA DNA LIGASE". Esta mezcla se incubó durante una noche a 12°C. Las siguientes etapas se realizaron según las instrucciones del kit TA Cloning[®] (British Biotechnology). Al final del procedimiento, las colonias blancas de bacterias recombinantes (blanco) se replantaron para ser cultivadas y la extracción de los plásmidos incorporados según el procedimiento denominado "miniprep" (17). La preparación de plásmido de cada colonia recombinante se escindió mediante una enzima de restricción apropiada y se analizó sobre gel de agarosa. Los plásmidos que tienen un inserto

detectado bajo luz UV después del marcado del gel con bromuro de etidio, se seleccionaron para la secuenciación del inserto, después de la hibridación con un cebador complementario del promotor Sp6 presente en el plásmido de clonación del TA Cloning kit®. Después, se ha efectuado la reacción previa a la secuenciación según el método preconizado para el uso del kit de secuenciación "Prism ready reaction kit dye deoxyterminator cycle sequencing kit" (Applied Biosystems, ref. 401384), y se ha realizado la secuenciación automática con el aparato "Automatic Sequencer, modelo 373 A Applied Biosystems según las instrucciones del fabricante.

Así, se ha podido obtener un clon denominado LTRGAG12 y se representa mediante su secuencia interna identificada mediante la SEC ID nº 60.

Este clon es probablemente representativo de elementos endógenos cercanos de ERV-9, presente en el ADN humano, en particular en el ADN de pacientes que padecen SEP, y que pueden interferir con la expresión del retrovirus MSRV-1, que por lo tanto, puede desempeñar un papel en la patogenia asociada al retrovirus MSRV-1 y que puede servir de marcador de una expresión específica en la patología implicada.

EJEMPLO 17: DETECCIÓN DE ANTICUERPOS ESPECÍFICOS ANTI-MSRV-1 EN EL SUERO HUMANO

La identificación de la secuencia del gen *pol* del retrovirus MSRV-1 y de una trama de lectura abierta de este gen ha permitido determinar la secuencia SEC ID nº 63 de aminoácidos de una región de dicho gen, referenciada mediante la SEC ID nº 62.

Se han ensayado diferentes péptidos sintéticos que corresponden a unos fragmentos de la secuencia proteica de la transcriptasa inversa MSRV-1 codificada por el gen *pol*, para su especificidad antigénica frente a sueros de pacientes que padecen SEP y de pacientes de control sanos.

Se han sintetizado los péptidos químicamente mediante síntesis en fase sólida, según la técnica de Merrifield (22). Las modalidades prácticas son las descritas a continuación.

a) Síntesis de los péptidos:

Se han sintetizado los péptidos sobre una resina de fenilacetamidometilo (PAM)/poliestireno/divinilbenceno (Applied Biosystems, Inc. Foster City, CA) usando un sintetizador automático "Applied Biosystems 430A". Los aminoácidos se enlazan en forma de ésteres de hidroxibenzotriazol (HOBt). Los aminoácidos usados proceden de Novabiochem (Laüflerlfingen, Suiza) o de Bachem (Bubendorf, Suiza).

La síntesis química se ha efectuado usando un protocolo de doble enlace con la N-metilpirrolidona (NMP) como disolvente. Los péptidos se han escindido de la resina así como las protecciones laterales, de manera simultánea, con la ayuda de ácido fluorhídrico (HF) en un aparato apropiado (aparato de escisión de tipo I, Peptide Institute, Osaka, Japón).

Para 1 g de peptidilresina, se usan 10 ml de HF, 1 ml de anisol y 1 ml de dimetilsulfuro 5DMS. La mezcla se agita durante 45 minutos a -2°C. Entonces, se evapora el HF al vacío. Después de lavados intensivos con éter, el péptido se eluye de la resina mediante ácido acético al 10%, y después se liofiliza.

Los péptidos se purifican mediante cromatografía líquida ultrarrápida preparativa sobre una columna VYDAC de tipo C18 (250 x 21 mm) (The Separation Group, Hesperia, CA, USA). La elución se realiza mediante un gradiente de acetonitrilo con un caudal de 22 ml/min. Las fracciones recogidas se controlan mediante una elución en condición isocrática sobre una columna VYDAC® C18 analítica (250 x 4,6 mm) con un caudal de 1 ml/min. Las fracciones que presentan el mismo tiempo de retención se reúnen y se liofilizan. Después, se analiza la fracción mayoritaria mediante cromatografía líquida ultrarrápida analítica, con el sistema descrito anteriormente. El péptido que se considera como de pureza aceptable se traduce mediante un pico único que representa 95% como mínimo del cromatograma.

Después, se analizan los péptidos purificados con el fin de controlar su composición de aminoácidos, con la ayuda de un analizador de aminoácidos automático. Applied Biosystems 420H. La medición de la masa molecular química (media) de los péptidos se obtiene usando la espectrometría de masas L.S.I.M.S en modo de ión positivo en un instrumento de doble focalización VG. ZAB.ZSEQ conectado a un sistema de adquisición DEC-VAX 2000 (VG analytical Ltd, Manchester, Inglaterra).

La reactividad de los diferentes péptidos se ha ensayado contra sueros de pacientes que padecen SEP y contra sueros de pacientes de control sanos. Esto ha permitido seleccionar un péptido denominado S24Q cuya secuencia se identifica mediante la SEC ID nº 63, codificado por una secuencia nucleotídica del gen *pol* de MSRV-1 (SEC ID nº 62).

b) Propiedades antigénicas:

Las propiedades antigénicas del péptido S24Q se han demostrado según el protocolo ELISA descrito a continuación.

- 5 El péptido S24Q liofilizado se ha disuelto en ácido acético al 10% con una concentración de 1 mg/ml. Esta disolución madre se ha alicuotado y se ha guardado a +4°C para el uso dentro de una quincena de días, o se ha congelado a -20°C, para un uso dentro de dos meses. Se ha diluido una alícuota en una disolución de PBS (Phosphate Buffer Saline) a fin de obtener una concentración final de péptido de 5 microgramos/ml. Se depositan 100 microlitros de esta dilución en cada pocillo de placas de microtitulación Nunc Maxisorb (nombre comercial). Las placas se recubren con un adhesivo de tipo "plate-sealer" y se mantienen durante 2 horas a +4°C, para la fase de adsorción del péptido sobre el plástico. El adhesivo se quita y las placas se lavan tres veces con un volumen de 300 microlitros de una disolución A (PBS 1x, 0,05% de Tween 20®), después se les da la vuelta sobre una tela absorbente. Las placas así escurridas se llenan con 250 microlitros por pocillo de una disolución B (disolución A + 10% de suero de cabra), y después se recubren con un adhesivo y se incuban durante 1 hora, a 37°C. Después, las placas se lavan tres veces con la disolución A, como se ha descrito anteriormente.

- Las muestras de suero a ensayar se diluyen previamente al 1/100 en la disolución B, y 100 microlitros de cada suero diluido a ensayar se depositan en los pocillos de cada placa de microtitulación. Se deposita un control negativo en un pocillo de cada placa, en forma de 100 microlitros de tampón B. Las placas recubiertas con un adhesivo se incuban entonces durante 1 hora 30 minutos, a 37°C. Después, las placas se lavan tres veces con la disolución A, como se ha descrito anteriormente. Para la respuesta IgG, se diluye un anticuerpo de cabra marcado con peroxidasa y dirigido contra los IgG humanos (comercializado por Jackson Immuno Research Inc.) en la disolución B (dilución 1/10.000). Entonces, se depositan 100 microlitros de la dilución adecuada del anticuerpo marcado en cada pocillo de las placas de microtitulación, y las placas recubiertas con un adhesivo se incuban durante 1 hora, a 37°C. Después, se efectúa nuevamente un lavado de las placas como se ha descrito anteriormente. Paralelamente, el sustrato de la peroxidasa se prepara según las indicaciones del kit bioMérieux. Se depositan 100 microlitros de disolución-sustrato en cada pocillo, y las placas se disponen al abrigo de la luz durante 20 a 30 minutos, a temperatura ambiente.

- Una vez la reacción colorada se ha estabilizado, se depositan 50 microlitros de Color 2 (nombre comercial de bioMérieux) en cada pocillo para detener la reacción. Las placas se disponen inmediatamente en un lector espectrofotométrico de placas ELISA, y la densidad óptica (D.O.) de cada pocillo se lee con una longitud de onda de 492 nm.

- Las muestras serológicas se depositan en doble o en triple, y se calcula la densidad óptica (D.O.) que corresponde al suero ensayado calculando la media de las D.O. obtenidas para una misma muestra, con la misma dilución.

La D.O. neta de cada suero corresponde a la D.O. media del suero de la que se resta la D.O. media del control negativo (disolución B: PBS, 0,05% de Tween 20®, 10% de suero de cabra).

40 c) Detección de anticuerpos IgG anti-MSRV-1 (S24Q) mediante ELISA:

Se ha usado la técnica descrita anteriormente con el péptido S24Q para buscar la presencia de anticuerpos IgG específicos anti-MSRV-1 en el suero de 15 pacientes, para los cuales se ha realizado un diagnóstico cierto de SEP según los criterios de Poser (23), y de 15 controles sanos (donantes de sangre).

- 45 En la figura 41, se representan los resultados de cada suero ensayado con un anticuerpo anti-IgG. Cada barra vertical representa la densidad óptica (D.O. a 492 nm) neta de un suero ensayado. El eje de ordenadas indica la D.O. neta en el pico de las barras verticales. Las 15 primeras barras verticales situadas a la izquierda de la línea discontinua vertical, representan los sueros de 15 controles sanos (donantes de sangre) y las 15 barras verticales situadas a la derecha de la línea discontinua vertical, representan los sueros de 15 casos de SEP ensayados. La gráfica permite visualizar 2 controles cuya D.O. emerge por encima de los valores agrupados de la población de control. Estos valores pueden representar la presencia de IgG específicos en pacientes seropositivos no-enfermos. Por lo tanto, se han evaluado dos métodos para determinar el umbral estadístico de positividad del ensayo.

- 55 La media de las D.O. netas de los controles, incluyendo los controles con D.O. netas elevadas, es de 0,129 y la desviación típica es de 0,06. Sin los 2 controles cuyas D.O. netas son superiores a 0,2, la media de los controles "negativos" es de 0,107 y la desviación típica de 0,03. Se puede calcular un umbral de positividad teórica según la fórmula:

- 60
$$\text{valor umbral} = (\text{media de las D.O. netas de los controles negativos}) + (2 \text{ ó } 3 \times \text{desviación típica de las D.O. netas de los controles negativos})$$

En el primer caso, se considera que existen seropositivos no-enfermos, y el valor umbral es igual a: $0,11 + (3 \times 0,03) = 0,20$. Los resultados negativos representan un "ruido de fondo" no específico de la presencia de anticuerpos específicamente dirigidos contra un epítipo del péptido.

65

En el segundo caso, si el conjunto de los controles constituidos por donantes de sangre con una buena salud aparente se coge como base de referencia, sin excluir los sueros *a priori* seropositivos, la desviación típica de los "controles no-SEP" es de 0,116. Entonces, el valor umbral es: $0,13 + (3 \times 0,006) = 0,31$.

Según este último análisis, el ensayo es específico de la SEP. En este aspecto, se constata que el ensayo es específico de la SEP, puesto que tal como se muestra en la tabla 1, ningún control tiene una D.O. neta por encima de este umbral. Realmente, este resultado refleja el hecho de que las titulaciones de los anticuerpos en los pacientes que padecen SEP son mayoritariamente más elevadas que en los controles sanos que han estado en contacto con MSRV-1.

En función del primer modo de cálculo y como se representa en la figura 41 y en la tabla 3, 6 de los 15 sueros SEP dan un resultado positivo (D.O. superior o igual a 0,2) indicando la presencia de IgG específicamente dirigidos contra el péptido S24Q, y por lo tanto contra una parte de la enzima de transcriptasa inversa del retrovirus MSRV-1 codificada por su gen pol y, por consiguiente, contra el retrovirus MSRV-1.

Así, aproximadamente 40% de los pacientes SEP ensayados han reaccionado contra un epítipo portado por el péptido S24Q y presentan IgG circulantes dirigidos contra este último.

De 15 donantes de sangre, aparentemente con una buena salud, 2 presentan un resultado positivo. Así, aparece que aproximadamente 13% de la población no-enferma puede haber estado en contacto de un epítipo portado por el péptido S24Q en condiciones que han llevado a una inmunización activa que se traduce mediante la persistencia de IgG séricos específicos. Estas condiciones son compatibles con una inmunización contra la transcriptasa inversa del retrovirus MSRV-1, durante una infección por el (y/o reactivación del) retrovirus MSRV-1. La ausencia de patología neurológica aparente que evoca la SEP en estos controles seropositivos puede indicar que son portadores sanos, que han eliminado un virus infeccioso después de haberse inmunizado o que constituyen una población con riesgo de portadores crónicos. En efecto, los datos epidemiológicos que muestran que un agente patógeno presente en el entorno de las regiones con alta preponderancia de SEP, puede ser la causa de esta enfermedad, implican que una fracción de la población libre de SEP ha estado necesariamente en contacto con dicho agente patógeno. Se ha demostrado que el retrovirus MSRV-1 constituye la totalidad o parte de este "agente patógeno" origen de la SEP, y es por lo tanto normal que unos controles cogidos en una población sana presenten anticuerpos de tipo IgG contra componentes del retrovirus MSRV-1.

Así, la detección de anticuerpos anti-S24Q en sólo una SEP de dos ensayados en este caso, puede reflejar el hecho de que este péptido no presenta ningún epítipo inmunodominante MSRV-1, que unas variaciones interindividuales de cepas pueden inducir una inmunización contra un motivo peptídico divergente en la misma región, o también que la evolución de la enfermedad y los tratamientos seguidos pueden modular en el tiempo la respuesta de anticuerpos contra el péptido S24Q.

TABLA nº 3

CONTROLES	SEP
0,101	0,136
0,058	0,391
0,126	0,37
0,131	0,119
0,105	0,267
0,294	0,141
0,116	0,102
0,088	0,18
0,105	0,411
0,172	0,164
0,137	0,049
0,223	0,644
0,08	0,268
0,073	0,065
0,132	0,074
Media	0,129
Desviación típica	0,06
Umbral	0,31

d) Detección de anticuerpos IgM anti-MSRV-1 mediante ELISA

Se ha usado la técnica de ELISA con el péptido S24Q para buscar la presencia de anticuerpos específicos IgM anti-MSRV-1 en los mismos sueros que anteriormente.

En la figura 42, se representan los resultados de cada suero ensayado con un anticuerpo anti-IgM. Cada barra vertical representa la densidad óptica (D.O. a 492 nm) neta de un suero ensayado. El eje de ordenadas indica la D.O. neta en el pico de las barras verticales. Las 15 primeras barras verticales situadas a la izquierda de la línea vertical que corta el eje de abscisas representan los sueros de 15 controles sanos (donantes de sangre), y las barras verticales situadas a la derecha de la línea discontinua vertical representan los sueros de 15 casos de SEP ensayados.

La media de las D.O. netas de las SEP ensayadas es de: 1,6.

La media de las D.O. netas de los controles es de: 0,7.

La desviación típica de los controles negativos es de: 0,6.

El umbral de positividad teórico se puede calcular según la fórmula:

$$\text{valor umbral} = (\text{media de las D.O. de los controles negativos}) + (3 \times \text{desviación típica de las D.O. de los controles negativos})$$

El valor umbral es, por lo tanto, igual a $0,7 + (3 \times 0,6) = 0,5$;

Los resultados negativos representan un "ruido de fondo" no específico de la presencia de anticuerpos específicamente dirigidos contra un epítipo del péptido.

Según este análisis y como se muestra en la figura 42 y en la tabla 4 correspondiente, el ensayo IgM es específico de la SEP, puesto que ningún control tiene una D.O. neta por encima del umbral. De los 15 sueros SEP 6 producen un resultado IgM positivo.

La diferencia de seroprevalencia entre los SEP y la población de control es extremadamente significativa: ensayo "chi-2", $p < 0,002$.

Estos resultados están a favor de un papel etiopatogénico de MSRV-1 en la SEP.

Así, la detección de los anticuerpos IgM e IgG contra el péptido S24Q permite evaluar, solo o en combinación con otros péptidos MSRV-1, la evolución de una infección mediante MSRV-1 y/o de la reactivación vírica de MSRV-1.

TABLA nº 4

CONTROLES	SEP
1,449	0,974
0,371	6,117
0,448	2,883
0,456	1,945
0,885	1,787
2,235	0,273
0,301	1,766
0,138	0,668
0,16	2,603
1,073	0,802
1,366	0,245
0,283	0,147
0,262	2,441
0,585	0,287
0,356	0,589
Media	0,7
Desviación típica	0,6
Umbral	2,5

Así, es posible, gracias a los descubrimientos efectuados y a los nuevos métodos realizados por los inventores, permitir la realización perfeccionada de los ensayos de diagnósticos de la infección y/o de la reactivación de MSRV-1, y de evaluar una terapia en la SEP y/o en la PR en base a su eficacia para "hacer negativa" la detección de estos agentes en los fluidos biológicos de los pacientes. Además, la detección precoz en personas que no presentan todavía señales neurológicas de SEP, o señales reumatológicas de PR, podría permitir instaurar un tratamiento tanto más eficaz de la evolución clínica posterior por cuanto que precedería el estado lesional que corresponde a la aparición de los trastornos clínicos. Ahora bien, hoy en día, no se puede establecer un diagnóstico de SEP o de PR antes de la instalación de una sintomatología lesional y, por lo tanto, no se instaura ningún tratamiento antes de la

emergencia de una clínica evocadora de lesiones ya consecuentes. El diagnóstico de una infección y/o de una reactivación de MSRV-1 y/o MSRV-2 en el ser humano es, por lo tanto, determinante, y la presente invención suministra los medios para ello.

- 5 Así, es posible, además de realizar un diagnóstico de la infección y/o de la reactivación de MSRV-1, evaluar una terapia en la SEP en base a su eficacia en "hacer negativa" la detección de estos agentes en los fluidos biológicos de los pacientes.

BIBLIOGRAFÍA

- 10 **(1)** Norrby E., Prog. Med. Virol., 1978; 24, 1-39.
- (2)** Johnson R.T., "Handbook of clinical neurology, 47 Demyelinating diseases", Vinken P. y Bruyn G.W., eds. Amsterdam, Elsevier Science Publishing, 1985, 319-336.
- 15 **(3)** Perron H. *et al.*, Res. Virol. 1989, 140, 551-561.
- (4)** Perron H. *et al.*, "Current concepts in multiple sclerosis" Wiethölter *et al.*, eds. Amsterdam, Elsevier, 1991, 111-116.
- 20 **(5)** Perron H. *et al.*, The Lancet 1991, 337, 862-863.
- (6)** Perron H. *et al.*, J. Gen. Virol. 1993, 74, 65-72.
- (7)** Fields y Knipe, Fundamental Virology 1986, Rev Press N.Y.
- 25 **(8)** Nielsen P.E. *et al.*, Science 1991; 254, 1497-1500.
- (9)** Maniatis *et al.*, Molecular Cloning, Cold Spring Harbor, 1982.
- 30 **(10)** Southern. E.M., J. Mol. Biol. 1975, 98, 503.
- (11)** Dunn A.R. y Hassel J.A., Cell 1977, 12, 23.
- (12)** Shih *et al.*, J. Virol. 1989, 63, 64-75.
- 35 **(13)** Perron H. *et al.*, Res. Vir. 1992, 143, 337-350.
- (14)** Meyerhans *et al.*, Cell 1989, 58, 901-910.
- 40 **(15)** Linial M.L. y Miller A.D., "Current topics in microbiology ADN immunobiology. Retroviruses, stratégies of replication" vol. 157, 125-152; Swanstrom R. y Vogt P.K., editores, Springer-Verlag, Heidelberg 1990.
- (16)** Lori F. *et al.*, J. Virol. 1992, 66, 5067-5074.
- 45 **(17)** Sambrook J., Fritsch E.F., y Maniatis T., Molecular cloning, a laboratory manual. Cold Spring Harbor Laboratory Press, 1989.
- (18)** La Mantia *et al.*, Nucleic Acids Research 1991, 19, 1513-1520.
- 50 **(19)** Gonzalez-Quintial R, Baccala R, Pope R M y Thoeofilopoulos N, J. Clin. Invest, vol. 97, Número 5, p. 1335-1343, 1996.
- (20)** Chomzynski P. y N. Sacchi, Analytical Biochemistry 1987, 162, 156-159.
- 55 **(21)** F. Mallet *et al.*, Journal of Clinical Microbiology 1993; 31, 1444-1449.
- (22)** G. Barany y R.B. Merrifield, 1980, en the Peptides, 2, 1-284, Gross E y Meienhofer J, Eds Academic Press, Nueva York.
- 60 **(23)** Poser *et al.*, Gbers G.C. eds. The diagnosis of multiple sclerosis Thieme Stratton Inc, Nueva York 1984: 225-229.
- (24)** La Mantia *et al.*, Nucleic Acid Research 1989, 17, 5913-22.
- 65 **(25)** PLAZA, A; KONO, D.H; THEOFILOPOULOS, A.N. NEW HUMAN V β GENES ADN POLYMORPHIC VARIANTS. J. Imm; 147(12): 4360-4365, 1991.

Listado de secuencias**(1) INFORMACIONES GENERALES:****(i) SOLICITANTE:**

- (A) NOMBRE: BIOMERIEUX
 (B) CALLE: NINGUNA
 (C) CIUDAD: MARCY L'ETOILE
 (E) PAÍS: FRANCIA
 (F) CÓDIGO POSTAL: 69280

- (ii) TÍTULO DE LA INVENCION: Polipéptidos antigénicos asociados a la esclerosis en placas y utilizaciones.
 (iii) NÚMERO DE SECUENCIAS: 92
 (iv) FORMATO LEGIBLE POR ORDENADOR:

- (A) TIPO DE SOPORTE: Disquete
 (B) ORDENADOR: IBM PC compatible
 (C) SISTEMA OPERATIVO: PC-DOS/MS-DOS
 (D) PROGRAMA: PatentIn Release #1.0, Versión #1.30 (OEB)

(2) INFORMACIONES PARA LA SEC ID nº: 1:**(i) CARACTERÍSTICAS DE LA SECUENCIA:**

- (A) LONGITUD: 1158 pares de bases
 (B) TIPO: Nucleótido
 (C) NÚMERO DE HEBRAS: simple
 (D) CONFIGURACIÓN: Lineal

- (ii) TIPO DE MOLÉCULA: ADNc
 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 1:

```

CCCTTTGCCA CTACATCAAT TTTAGGAGTA AGGAAACCCA ACGGACAGTG GAGGTTAGTG 60
CAAGAACTCA GGATTATCAA TGAGGCTGTT GTTCCTCTAT ACCCAGCTGT ACCTAACCCCT 120
TATACAGTGC TTTCCCAAAT ACCAGAGGAA GCAGAGTGGT TTACAGTCCT GGACCTTAAG 180
GATGCCTTT TCTGCATCCC TGTACGTCCT GACTCTCAAT TCTTGTTTGC CTTTGAAGAT 240
CCTTTGAACC CAACGTCTCA ACTCACCTGG ACTGTTTTAC CCCAAGGGTT CAGGGATAGC 300
CCCCATCTAT TTGGCCAGGC ATTAGCCCAA GACTTGAGTC AATTCTCATA CCTGGACACT 360
CTTGTCCTTC AGTACATGGA TGATTTACTT TTAGTCGCCC GTTCAGAAAC CTTGTGCCAT 420
CAAGCCACCC AAGAACTCTT AACTTTCTCT ACTACCTGTG GCTACAAGGT TTCCAAACCA 480
AAGGCTCGGC TCTGCTCACA GGAGATTAGA TACTNAGGGC TAAAATTATC CAAAGGCACC 540
AGGGCCCTCA GTGAGGAACG TATCCAGCCT ATACTGGCTT ATCCTCATCC CAAAACCCTA 600
AAGCAACTAA GAGGGTTTCT TGGCATAACA GGTTCCTGCC GAAAACAGAT TCCCAGGTAC 660
ASCCCAATAG CCAGACCATT ATATACACTA ATTANGGAAA CTCAGAAAGC CAATACCTAT 720
TTAGTAAGAT GGACACCCTAC AGAAGTGGCT TTCCAGGCCC TAAAGAAGGC CCTAACCCAA 780
GCCCCAGTGT TCAGCTTGCC AACAGGGCAA GATTTTTCTT TATATGCCAC AGAAAAACA 840
GGAATAGCTC TAGGAGTCCT TACGCAGGTC TCAGGGATGA GCTTGCAACC CGTGGTATAC 900
CTGAGTAAGG AAATTGATGT AGTGGCAAAG GGTGGCCTC ATNGTTTATG GGTAAATGGNG 960
GCAGTAGCAG TCTNAGTATC TGAAGCAGTT AAAATAATAC AGGGAAGAGA TCTTNCTGTG 1020
TGGACATCTC ATGATGTGAA CGGCATACTC ACTGCTAAAG GAGACTTGTG GTTGTGACAG 1080
AACCATTTAC TTAANTATCA GGCTCTATTA CTTGAAGAGC CAGTGCTGNG ACTGCGCACT 1140
TGTGCAACTC TTAAACCC 1158

```

(2) INFORMACIONES PARA LA SEC ID nº: 2:**(i) CARACTERÍSTICAS DE LA SECUENCIA:**

- (A) LONGITUD: 297 pares de bases
 (B) TIPO: Nucleótido
 (C) NÚMERO DE HEBRAS: simple
 (D) CONFIGURACIÓN: Lineal

- (ii) TIPO DE MOLÉCULA: ADNc

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 2:

CCCTTTGCCA	CTACATCAAT	TTTAGGAGTA	AGGAAACCCA	ACGGACAGTG	GAGGTTAGTG	60
CAAGAACTCA	GGATTATCAA	TGAGGCTGTT	GTTCTCTAT	ACCCAGCTGT	ACCTAACCCT	120
TATACAGTGC	TTTCCCAAAT	ACCAGAGGAA	GCAGAGTGGT	TTACAGTCCT	GGACCTTAAG	180
GATGCCCTTT	TCTGCATCCC	TGTACGTCCT	GA CTCTCAAT	TCTTGTTTGC	CTTTGAAGAT	240
CCTTTGAACC	CAACGTCTCA	ACTCACCTGG	ACTGTTTTAC	CCCAAGGGTT	CAAGGGA	297

5 (2) INFORMACIONES PARA LA SEC ID nº: 3:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- 10 (A) LONGITUD: 85 pares de bases
(B) TIPO: Nucleótido
(C) NÚMERO DE HEBRAS: simple
(D) CONFIGURACIÓN: Lineal

(ii) TIPO DE MOLÉCULA: ADNc

15 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 3:

GTTTAGGGAT	ANCCCTCATC	TCTTTGGTCA	GGTACTGGCC	CAAGATCTAG	GCCACTTCTC	60
AGGTCCAGSN	ACTCTGTYCC	TTCAG				85

(2) INFORMACIONES PARA LA SEC ID nº: 4:

20

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- 25 (A) LONGITUD: 86 pares de bases
(B) TIPO: Nucleótido
(C) NÚMERO DE HEBRAS: simple
(D) CONFIGURACIÓN: Lineal

(ii) TIPO DE MOLÉCULA: ADNc

30 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 4:

GTTTCAGGGAT	AGCCCCCATC	TATTTGGCCA	GGCACTAGCT	CAATACTTGA	GCCAGTTCTC	60
ATACCTGGAC	AYTCTYGTCC	TTCGGT				86

(2) INFORMACIONES PARA LA SEC ID nº: 5:

35 (i) CARACTERÍSTICAS DE LA SECUENCIA:

- 40 (A) LONGITUD: 85 pares de bases
(B) TIPO: Nucleótido
(C) NÚMERO DE HEBRAS: simple
(D) CONFIGURACIÓN: Lineal

(ii) TIPO DE MOLÉCULA: ADNc

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 5:

GTTCARRGA	TAGCCCCCATC	TATTTGGCCW	RGYATTAGCC	CAAGACTTGA	GYCAATTCTC	60
ATACCTGGA	CACTCTGTCC	TTYRG				85

45

(2) INFORMACIONES PARA LA SEC ID nº: 6:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

50

- (A) LONGITUD: 85 pares de bases
(B) TIPO: Nucleótido
(C) NÚMERO DE HEBRAS: simple
(D) CONFIGURACIÓN: Lineal

55

(ii) TIPO DE MOLÉCULA: ADNc

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 6:

GTTTCAGGGAT AGCTCCCATC TATTTGGCCT GGCATTAACC CGAGACTTAA GCCAGTTCTY 60
ATACGTGGAC ACTCTTGTCC TTTGG 85

5 (2) INFORMACIONES PARA LA SEC ID nº: 7:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- 10 (A) LONGITUD: 111 pares de bases
 (B) TIPO: Nucleótido
 (C) NÚMERO DE HEBRAS: simple
 (D) CONFIGURACIÓN: Lineal

(ii) TIPO DE MOLÉCULA: ADNc

15 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 7:

GTGTTGCCAC AGGGGTTTAR RGATANCYCY CATCTMTTTG GYCWRGYAYT RRCYCRAKAY 60
YTRRGYCAVT TCTYAKRYSY RGSNAYTCTB KYCCTTYRGT ACATGGATGA C 111

(2) INFORMACIONES PARA LA SEC ID nº: 8:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- 20 (A) LONGITUD: 645 pares de bases
 (B) TIPO: Nucleótido
 25 (C) NÚMERO DE HEBRAS: simple
 (D) CONFIGURACIÓN: Lineal

(ii) TIPO DE MOLÉCULA: ADNc

30 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 8:

TCAGGGATAG CCCCCATCTA TTTGGCCAGG CATTAGCCCA AGACTTGAGT CAATTCTCAT 60
ACCTGGACAC TCTTGTCCTT CAGTACATGG ATGATTTACT TTTAGTCGCC CGTTCAGAAA 120
CCTTGTCGCA TCAAGCCACC CAAGAACTCT TAACTTTCTT CACTACCTGT GGCTACAAGG 180
TTTCCAAACC AAAGGCTCGG CTCTGCTCAC AGGAGATTAG ATACTNAGGG CTAAAATTAT 240
CCAAAGGCAC CAGGGCCCTC AGTGAGGAAC GTATCCAGCC TATACTGGCT TATCCTCATC 300
CCAAAACCTT AAAGCAACTA AGAGGGTTCC TTGGCATAAC AGGTTTCTGC CGAAAACAGA 360
TTCCAGGTA CASCCTAATA GCCAGACCAT TATATACACT AATTANGGAA ACTCAGAAAG 420
CCAATACCTA TTTAGTAAGA TGGACACCTA CAGAAGTGGC TTTCCAGGCC CTAAAGAAGG 480
CCCTAACCCA AGCCCCAGTG TTCAGCTTGC CAACAGGGCA AGATTTTCTT TTATATGCCA 540
CAGAAAAAAC AGGAATAGCT CTAGGAGTCC TTACGCAGGT CTCAGGGATG AGCTTGCAAC 600
CCGTGGTATA CCTGAGTAAG GAAATTGATG TAGTGCCAAA GGGTT 645

(2) INFORMACIONES PARA LA SEC ID nº: 9:

35 (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 741 pares de bases
 (B) TIPO: Nucleótido
 40 (C) NÚMERO DE HEBRAS: simple
 (D) CONFIGURACIÓN: Lineal

(ii) TIPO DE MOLÉCULA: ADNc

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 9:

CAAGCCACCC	AAGAACTCTT	AAATTTCTCTC	ACTACCTGTG	GCTACAAGGT	TTCCAAACCA	60
AAGGCTCAGC	TCTGCTCACA	GGAGATTAGA	TACTTAGGGT	TAAAATTATC	CAAAGGCACC	120
AGGGGCCTCA	GTGAGGAACG	TATCCAGCCT	ATACTGGGT	ATCCTCATCC	CAAACCCCTA	180
AAGCAACTAA	GAGGGTTCCT	TAGCATGATC	AGGTTTCTGC	CGAAAACAAG	ATTCCCAGGT	240
ACAACCAAAA	TAGCCAGACC	ATTATATACA	CTAATTAAGG	AAACTCAGAA	AGCCAATACC	300
TATTTAGTAA	GATGGACACC	TAAACAGAAG	GCTTTCCAGG	CCCTAAAGAA	GGCCCTAACC	360
CAAGCCCCAG	TGTTTCAGCTT	GCCAACAGGG	CAAGATTTTT	CTTTATATGG	CACAGAAAAA	420
ACAGGAATCG	CTCTAGGAGT	CCTTACACAG	GTCCGAGGGA	TGAGCTTGCA	ACCCGTGGCA	480
TACCTGAATA	AGGAAATTGA	TGTAGTGGCA	AAGGGTTGGC	CTCATNGTTT	ATGGGTAAATG	540
GNGGCAGTAG	CAGTCTNAGT	ATCTGAAGCA	GTTAAATAA	TACAGGGAAG	AGATCTTNCT	600
GTGTGGACAT	CTCATGATGT	GAACGGCATA	CTCACTGCTA	AAGGAGACTT	GTGGTTGTCA	660
GACAACCATT	TACTTAANTA	TCAGGCTCTA	TTACTTGAAG	AGCCAGTGCT	GNGACTGCGC	720
ACTTGTGCAA	CTCTTAAACC	C				741

(2) INFORMACIONES PARA LA SEC ID nº: 10:

5 (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 93 pares de bases
- (B) TIPO: Nucleótido
- (C) NÚMERO DE HEBRAS: simple
- (D) CONFIGURACIÓN: Lineal

10

- (ii) TIPO DE MOLÉCULA: ADNc
- (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 10:

TGGAAAGTGT	TGCCACAGGG	CGCTGAAGCC	TATCGCGTGC	AGTTGCCGGA	TGCCGCCTAT	60
AGCCTCTACA	TGGATGACAT	CCTGCTGGCC	TCC			93

15

(2) INFORMACIONES PARA LA SEC ID nº: 11:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 96 pares de bases
- (B) TIPO: Nucleótido
- (C) NÚMERO DE HEBRAS: simple
- (D) CONFIGURACIÓN: Lineal

20

- (ii) TIPO DE MOLÉCULA: ADNc
- (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 11:

TTGGATCCAG	TGYTGCCACA	GGGCGCTGAA	GCCTATCGCG	TGCAGTTGCC	GGATGCCGCC	60
TATAGCCTCT	ACGTGGATGA	CCTSCTGAAG	CTTGAG			96

30

(2) INFORMACIONES PARA LA SEC ID nº: 12:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 748 pares de bases
- (B) TIPO: Nucleótido
- (C) NÚMERO DE HEBRAS: simple
- (D) CONFIGURACIÓN: Lineal

35

- (ii) TIPO DE MOLÉCULA: ADNc
- (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 12:

40

TGCAAGCTTC	ACCGCTTGCT	GGATGTAGGC	CTCAGTACCG	GNGTGCCCCG	CGCGCTGTAG	60
TTCGATGTAG	AAAGCGCCCG	GAAACACGCG	GGACCAATGC	GTCGCCAGCT	TGCGCGCCAG	120
CGCCTCGTTG	CCATTGGCCA	GCGCCACGCC	GATATCACCC	GCCATGGCGC	CGGAGAGCGC	180
CAGCAGACCG	GCGGCCAGCG	GCGCATTCTC	AACGCCGGGC	TCGTCTGAACC	ATTCGGGGGC	240
GATTTCGCA	CGACCGCGAT	GCTGGTTGGA	GAGCCAGGCC	CTGGCCAGCA	ACTGGCACAG	300
GTTCAGGTAA	CCCTGCTTGT	CCCGCACCAA	CAGCAGCAGG	CGGGTCGGCT	TGTCGCGCTC	360
GTCGTGATTG	GTGATCCACA	CGTCAGCCCC	GACGATGGGC	TTCACGCCCT	TGCCACGCGC	420
TTCCTTGTAG	ANGCGACCA	GCCCCAAGGC	ATTGGCGAGA	TCGGTCAGCG	CCAAGGCGCC	480
CATGCCATCT	TTGGCGGCAG	CCTTGACGGC	ATCGTCGAGA	CGGACATTGC	CATCGACGAC	540
GGAATATTCT	GAGTGGAGAC	GGAGGTGGAC	GAAGCGCGGC	GAATTCATCC	GCGTATTGTA	600
ACGGGTGACA	CCTTCCGCAA	AGCATTCCGG	ACGTGCCCCG	TTGACCCGGA	GCAACCCCGC	660
ACGGCTGCGC	GGGCAGTTAT	AATTTCCGGT	TACGAATCAA	CGGGTTACCC	CAGGGCGCTG	720
AAGCCTATCG	CGTGCAGTTG	CCGGATGC				748

(2) INFORMACIONES PARA LA SEC ID nº: 13:

5 (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 18 pares de bases
- (B) TIPO: Nucleótido
- (C) NÚMERO DE HEBRAS: simple
- (D) CONFIGURACIÓN: Lineal

(ii) TIPO DE MOLÉCULA: ADNc

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 13:

15 GCATCCGGCA ACTGCACG

18

(2) INFORMACIONES PARA LA SEC ID nº: 14:

20 (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 20 pares de bases
- (B) TIPO: Nucleótido
- (C) NÚMERO DE HEBRAS: simple
- (D) CONFIGURACIÓN: Lineal

(ii) TIPO DE MOLÉCULA: ADNc

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 14:

30 GTAGTTCGAT GTAGAAAGCG

20

(2) INFORMACIONES PARA LA SEC ID nº: 15:

35 (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 18 pares de bases
- (B) TIPO: Nucleótido
- (C) NÚMERO DE HEBRAS: simple
- (D) CONFIGURACIÓN: Lineal

(ii) TIPO DE MOLÉCULA: ADNc

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 15:

45 GCATCCGGCA ACTGCACG

18

(2) INFORMACIONES PARA LA SEC ID nº: 16:

50 (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 23 pares de bases
- (B) TIPO: Nucleótido
- (C) NÚMERO DE HEBRAS: simple
- (D) CONFIGURACIÓN: Lineal

(ii) TIPO DE MOLÉCULA: ADNc

	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 16:	
	AGGAGTAAGG AAACCCAACG GAC	23
5	(2) INFORMACIONES PARA LA SEC ID nº: 17:	
	(i) CARACTERÍSTICAS DE LA SECUENCIA:	
10	(A) LONGITUD: 19 pares de bases	
	(B) TIPO: Nucleótido	
	(C) NÚMERO DE HEBRAS: simple	
	(D) CONFIGURACIÓN: Lineal	
	(ii) TIPO DE MOLÉCULA: ADNc	
15	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 17:	
	TAAGAGTTGC ACAAGTGCG	19
20	(2) INFORMACIONES PARA LA SEC ID nº: 18:	
	(i) CARACTERÍSTICAS DE LA SECUENCIA:	
25	(A) LONGITUD: 21 pares de bases	
	(B) TIPO: Nucleótido	
	(C) NÚMERO DE HEBRAS: simple	
	(D) CONFIGURACIÓN: Lineal	
	(ii) TIPO DE MOLÉCULA: ADNc	
30	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 18:	
	TCAGGGATAG CCCCCATCTA T	21
	(2) INFORMACIONES PARA LA SEC ID nº: 19:	
35	(i) CARACTERÍSTICAS DE LA SECUENCIA:	
	(A) LONGITUD: 24 pares de bases	
	(B) TIPO: Nucleótido	
40	(C) NÚMERO DE HEBRAS: simple	
	(D) CONFIGURACIÓN: Lineal	
	(ii) TIPO DE MOLÉCULA: ADNc	
	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 19:	
45	AACCCTTTGC CACTACATCA ATTT	24
	(2) INFORMACIONES PARA LA SEC ID nº: 20:	
50	(i) CARACTERÍSTICAS DE LA SECUENCIA:	
	(A) LONGITUD: 15 pares de bases	
	(B) TIPO: Nucleótido	
	(C) NÚMERO DE HEBRAS: simple	
55	(D) CONFIGURACIÓN: Lineal	
	(ii) TIPO DE MOLÉCULA: ADNc	
	(ix) CARACTERÍSTICAS:	
60	(B) EMPLAZAMIENTOS: 5, 7, 10, 13	
	(D) OTRAS INFORMACIONES: G representa la insosina (i)	
	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 20:	
65	GGTCGTGCCG CAGGG	15
	(2) INFORMACIONES PARA LA SEC ID nº: 21:	

	(i) CARACTERÍSTICAS DE LA SECUENCIA:	
5	(A) LONGITUD: 21 pares de bases (B) TIPO: Nucleótido (C) NÚMERO DE HEBRAS: simple (D) CONFIGURACIÓN: Lineal	
10	(ii) TIPO DE MOLÉCULA: ADNc (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 21:	
	TTAGGGATAG CCCTCATCTC T	21
	(2) INFORMACIONES PARA LA SEC ID nº: 22:	
15	(i) CARACTERÍSTICAS DE LA SECUENCIA:	
20	(A) LONGITUD: 21 pares de bases (B) TIPO: Nucleótido (C) NÚMERO DE HEBRAS: simple (D) CONFIGURACIÓN: Lineal	
25	(ii) TIPO DE MOLÉCULA: ADNc (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 22:	
	TCAGGGATAG CCCCCATCTA T	21
	(2) INFORMACIONES PARA LA SEC ID nº: 23:	
30	(i) CARACTERÍSTICAS DE LA SECUENCIA:	
35	(A) LONGITUD: 24 pares de bases (B) TIPO: Nucleótido (C) NÚMERO DE HEBRAS: simple (D) CONFIGURACIÓN: Lineal	
40	(ii) TIPO DE MOLÉCULA: ADNc (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 23:	
	AACCCTTTGC CACTACATCA ATTT	24
	(2) INFORMACIONES PARA LA SEC ID nº: 24:	
45	(i) CARACTERÍSTICAS DE LA SECUENCIA:	
50	(A) LONGITUD: 23 pares de bases (B) TIPO: Nucleótido (C) NÚMERO DE HEBRAS: simple (D) CONFIGURACIÓN: Lineal	
55	(ii) TIPO DE MOLÉCULA: ADNc (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 24:	
	GCGTAAGGAC TCCTAGAGCT ATT	23
	(2) INFORMACIONES PARA LA SEC ID nº: 25:	
60	(i) CARACTERÍSTICAS DE LA SECUENCIA:	
65	(A) LONGITUD: 18 pares de bases (B) TIPO: Nucleótido (C) NÚMERO DE HEBRAS: simple (D) CONFIGURACIÓN: Lineal	
	(ii) TIPO DE MOLÉCULA: ADNc (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 25:	

	TCATCCATGT ACCGAAGG	18
5	(2) INFORMACIONES PARA LA SEC ID nº: 26:	
	(i) CARACTERÍSTICAS DE LA SECUENCIA:	
	(A) LONGITUD: 20 pares de bases	
10	(B) TIPO: Nucleótido	
	(C) NÚMERO DE HEBRAS: simple	
	(D) CONFIGURACIÓN: Lineal	
	(ii) TIPO DE MOLÉCULA: ADNc	
15	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 26:	
	ATGGGGTTCC CAAGTTCCT	20
	(2) INFORMACIONES PARA LA SEC ID nº: 27:	
20	(i) CARACTERÍSTICAS DE LA SECUENCIA:	
	(A) LONGITUD: 20 pares de bases	
	(B) TIPO: Nucleótido	
25	(C) NÚMERO DE HEBRAS: simple	
	(D) CONFIGURACIÓN: Lineal	
	(ii) TIPO DE MOLÉCULA: ADNc	
	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 27:	
30	GCCGATATCA CCCGCCATGG	20
	(2) INFORMACIONES PARA LA SEC ID nº: 28:	
35	(i) CARACTERÍSTICAS DE LA SECUENCIA:	
	(A) LONGITUD: 18 pares de bases	
	(B) TIPO: Nucleótido	
	(C) NÚMERO DE HEBRAS: simple	
40	(D) CONFIGURACIÓN: Lineal	
	(ii) TIPO DE MOLÉCULA: ADNc	
	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 28:	
45	GCATCCGGCA ACTGCACG	18
	(2) INFORMACIONES PARA LA SEC ID nº: 29:	
	(i) CARACTERÍSTICAS DE LA SECUENCIA:	
50	(A) LONGITUD: 20 pares de bases	
	(B) TIPO: Nucleótido	
	(C) NÚMERO DE HEBRAS: simple	
	(D) CONFIGURACIÓN: Lineal	
55	(ii) TIPO DE MOLÉCULA: ADNc	
	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 29:	
	CGCGATGCTG GTTGGAGAGC	20
60	(2) INFORMACIONES PARA LA SEC ID nº: 30:	
	(i) CARACTERÍSTICAS DE LA SECUENCIA:	
65	(A) LONGITUD: 20 pares de bases	
	(B) TIPO: Nucleótido	
	(C) NÚMERO DE HEBRAS: simple	

	(D) CONFIGURACIÓN: Lineal	
	(ii) TIPO DE MOLÉCULA: ADNc	
5	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 30:	
	TCTCCACTCC GAATATTCCG	20
	(2) INFORMACIONES PARA LA SEC ID nº: 31:	
10	(i) CARACTERÍSTICAS DE LA SECUENCIA:	
	(A) LONGITUD: 26 pares de bases	
	(B) TIPO: Nucleótido	
	(C) NÚMERO DE HEBRAS: simple	
15	(D) CONFIGURACIÓN: Lineal	
	(ii) TIPO DE MOLÉCULA: ADNc	
	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 31:	
20	GATCTAGGCC ACTTCTCAGG TCCAGS	26
	(2) INFORMACIONES PARA LA SEC ID nº: 32:	
	(i) CARACTERÍSTICAS DE LA SECUENCIA:	
25	(A) LONGITUD: 23 pares de bases	
	(B) TIPO: Nucleótido	
	(C) NÚMERO DE HEBRAS: simple	
	(D) CONFIGURACIÓN: Lineal	
30	(ii) TIPO DE MOLÉCULA: ADNc	
	(ix) CARACTERÍSTICAS:	
	(B) EMPLAZAMIENTO: 6, 12, 19	
35	(D) OTRAS INFORMACIONES: G representa la inosina (i)	
	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 32:	
	CATCTGTTTG GGCAGGCAGT AGC	23
40	2) INFORMACIONES PARA LA SEC ID nº: 33:	
	(i) CARACTERÍSTICAS DE LA SECUENCIA:	
45	(A) LONGITUD: 24 pares de bases	
	(B) TIPO: Nucleótido	
	(C) NÚMERO DE HEBRAS: simple	
	(D) CONFIGURACIÓN: Lineal	
50	(ii) TIPO DE MOLÉCULA: ADNc	
	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 33:	
	CTTGAGCCAG TTCTCATACC TGGA	24
55	2) INFORMACIONES PARA LA SEC ID nº: 34:	
	(i) CARACTERÍSTICAS DE LA SECUENCIA:	
	(A) LONGITUD: 22 pares de bases	
60	(B) TIPO: Nucleótido	
	(C) NÚMERO DE HEBRAS: simple	
	(D) CONFIGURACIÓN: Lineal	
	(ii) TIPO DE MOLÉCULA: ADNc	
65	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 34:	

AGTGYTRCCM CARGGCGCTG AA

22

2) INFORMACIONES PARA LA SEC ID nº: 35:

5 (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 22 pares de bases

(B) TIPO: Nucleótido

(C) NÚMERO DE HEBRAS: simple

10 (D) CONFIGURACIÓN: Lineal

(ii) TIPO DE MOLÉCULA: ADNc

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 35:

15 GMGGCCAGCA GSAKGTCATC CA

22

2) INFORMACIONES PARA LA SEC ID nº: 36:

20 (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 22 pares de bases

(B) TIPO: Nucleótido

(C) NÚMERO DE HEBRAS: simple

25 (D) CONFIGURACIÓN: Lineal

(ii) TIPO DE MOLÉCULA: ADNc

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 36:

30 GGATGCCGCC TATAGCCTCT AC

22

2) INFORMACIONES PARA LA SEC ID nº: 37:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

35 (A) LONGITUD: 22 pares de bases

(B) TIPO: Nucleótido

(C) NÚMERO DE HEBRAS: simple

(D) CONFIGURACIÓN: Lineal

40 (ii) TIPO DE MOLÉCULA: ADNc

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 37:

AAGCCTATCG CGTGCA GTTG CC

22

45 (2) INFORMACIONES PARA LA SEC ID nº: 38:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 40 pares de bases

50 (B) TIPO: Nucleótido

(C) NÚMERO DE HEBRAS: simple

(D) CONFIGURACIÓN: Lineal

(ii) TIPO DE MOLÉCULA: ADNc

55 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 38:

TAAAGATCTA GAATTCGGCT ATAGGCGGCA TCCGGCAAGT

40

(2) INFORMACIONES PARA LA SEC ID nº: 39:

60 (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 50 aminoácidos

(B) TIPO: aminoácidos

65 (ii) TIPO DE MOLÉCULA: péptido

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 39:

```

    Asp Ala Phe Phe Cys Ile Pro Val Arg Pro Asp Ser Gln Phe Leu Phe
      1      5      10
    Ala Phe Glu Asp Pro Leu Asn Pro Thr Ser Gln Leu Thr Trp Thr Val
      20      25      30
    Leu Pro Gln Gly Phe Arg Asp Ser Pro His Leu Phe Gly Gln Ala Leu
      35      40      45
    Ala Gln
      50
  
```

5 (2) INFORMACIONES PARA LA SEC ID nº: 40:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- 10 (A) LONGITUD: 150 pares de bases
 (B) TIPO: Nucleótido
 (C) NÚMERO DE HEBRAS: simple
 (D) CONFIGURACIÓN: Lineal

(ii) TIPO DE MOLÉCULA: ADNc

15 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 40:

```

GATGCCTTTT TCTGCATCCC TGTACGTCCT GACTCTCAAT TCTTGTTTGC CTTTGAAGAT 60
CCTTTGAACC CAACGTCTCA ACTCACCTGG ACTGTTTTAC CCCAAGGGTT CAGGGATAGC 120
CCCCATCTAT TTGGCCAGGC ATTAGCCCAA 150
  
```

(2) INFORMACIONES PARA LA SEC ID nº: 41:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- 20 (A) LONGITUD: 11 aminoácidos
 (B) TIPO: aminoácidos

25 (ii) TIPO DE MOLÉCULA: péptido

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 41:

```

Cys Ile Pro Val Arg Pro Asp Ser Gln Phe Leu
  1      5      10
  
```

(2) INFORMACIONES PARA LA SEC ID nº: 42:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- 35 (A) LONGITUD: 17 aminoácidos
 (B) TIPO: aminoácidos

(ii) TIPO DE MOLÉCULA: péptido

40 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 42:

```

Val Leu Pro Gln Gly Phe Arg Asp Ser Pro His Leu Phe Gly Glu Ala
  1      5      10      15
Leu
17
  
```

(2) INFORMACIONES PARA LA SEC ID nº: 43:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- 45 (A) LONGITUD: 8 aminoácidos
 (B) TIPO: aminoácidos

50 (ii) TIPO DE MOLÉCULA: péptido

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 43:

Leu Phe Ala Phe Glu Asp Pro Leu
1 5 8

5 (2) INFORMACIONES PARA LA SEC ID nº: 44:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- 10 (A) LONGITUD: 8 aminoácidos
 (B) TIPO: aminoácidos

(ii) TIPO DE MOLÉCULA: péptido
 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 44:

15 **Phe Ala Phe Glu Asp Pro Leu Asn**
1 5 8

(2) INFORMACIONES PARA LA SEC ID nº: 45:

20 (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 25 pares de bases
 (B) TIPO: Nucleótido
 (C) NÚMERO DE HEBRAS: simple
 (D) CONFIGURACIÓN: Lineal

25 (ii) TIPO DE MOLÉCULA: ADNc
 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 45:

30 GTGCTGATTG GTGTATTTAC AATCC

25

(2) INFORMACIONES PARA LA SEC ID nº: 46:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- 35 (A) LONGITUD: 1859 pares de bases
 (B) TIPO: Nucleótido
 (C) NÚMERO DE HEBRAS: simple
 (D) CONFIGURACIÓN: Lineal

40 (ii) TIPO DE MOLÉCULA: ADNc
 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 46:

```

GTGCTGATTG GTGTATTTAC AATCCTTTAT CTAATCCGAA ATGCCCATGT TGCAATATGG 60
AAAGAAAGGG AGTTCCTAAC CTCTGGGGGA ACCCCCATTA AATACCACAA GTAAATCATG 120
GAGTTATTGC ACACAGTGCA AAAACTCAAG GAGGTGGAAG TCTTACACTG CCAAAGCCAT 180
CAGAAAAGGG AAGAGGGGAG AAGAGCAGCA TAAGTGCGTA CAGAGGCAAG GAAAGACTAG 240
CAGAAAGGAA AGAGAGAAAG AGACAGAAAG TCAGAGAGAG AGAGAGGAAG AGACAGAGCA 300
CAAAGAGGGA GTCAGAGAGA GAGAGAGACA GAGAGTCAGA GAGAAGGAAA GAGAGAGAGG 360
AAGAGACAAA GAATGAATCA AACAGAGAGA CAGAAAGTCA GAGAGAGAGA GAGAGAGGAA 420
GAGACAGAGA AAAAGAGGGA GTCAGAAAAA GAGAGACCAA AGAAGAAGTC CAAAGAGAAA 480
GAAAGAGAGA TGAAGTAGT AAAGGAAAAA CAGTGTACCC TATTCCTTTA AAAGCCGGGG 540
TAAATTTAAA ACCTATAATT GATAACTGAA GGTCTTCTCT GTAACCCTGT AACACTCCAA 600
TACCACCTTG TTGTCAAGTG TAAACAAGGG CGTAGCCCAA AAGCACTGAG GCCACTAACA 660
ACCCATAGCC TTCCTATCAA AATTCCTTAA CCCAGCAGGT TTCCTAACAG GGGATCTAAA 720
TCTTAATTAA TTACCATAA ATGGTCCAAC CAGACTTAGG AGGAATTCCC TTCAGGACGG 780
GAAGATAGAT GCTTCCTCCC AGGCGATTAA GGGAGAAAGA CACAATGGGT ATTCAAGTAAG 840
TGCCAAGGGG AACACTTGTA GAAGCAAAGT TAGGAAAATT GCCAAATAAT TGGTTTGCTC 900
AAGAGTTGTT TGCACCTCAGC CAAACCTTGA AGTACTTGCA GAATCAGAAA GGAGCCATCT 960
ATACCAATTC TAAGTTAATA TGGACTGAAG GAGGTTTTAT TAATACCAA GAGAAATTAA 1020
AATCCCAAAC TTATAAGGTT TTCAACCAA GTAAAGTTTG CTAAAAGTTA ACAGCGTAAC 1080
ATGTATTATC CTAATACCAC AACTCTCAA AGGATTTCTC AGACAGTTTG CAAGAAATAA 1140
TGATATCTAT CCTTACTCTA CAATCCCAA TAGACTCTTT GGCAGCAGTG ACTCTCCAAA 1200
ACCGTCAAGG CCTAGACCTC CTCACTGCTG AGAAAGGAGG ACTCTGCACC TTCTTAAGGG 1260
AAGAGTGTTG TCTTTACACT AACCAGTCAG GGATAGTATG AGATGCTGCC CGGCATTTAC 1320
AGAAAAAGGC TTCTGAAATC AGACAACGCC TTTCAAATTC CTATACCAAC CTCTGGAGTT 1380
GGGCAACATG GTTCTTTCCC TTTCTATGTC CCATGGCTGC CATCTTGCTA TTAATCGCCT 1440
TTGGGCCCTG TATTTTTAAC CTCCTTGTC AATTTGTTTC TTCTAGGATC GAGGCCATCA 1500
AGCTACAGAT GGTCTTACAA ATGGAACCCC AAATGAGCTC AACTATCAAC TTCTACTGAG 1560
GACCCCTAGA CCAACCCCTT GGCCTTTTCA CTGGCCTAAA GAGTTCCCCT CTGGAGGACA 1620
CTACCACTGC AGGGCCCCAT CTTTGCCCT ATCCAGAAGG AAGTAGCTAG AGCAGTCATT 1680
GCCCAATTCC CAAGAGCAGC TGGGGTGTC CGTTTAGAGT GGGGATTGAG AGGTGAAGCC 1740
AGCTGGACTT CTGGGTCGGG TGGGGACTTG GAGAACTTTT GTGTCTAGCT AAAGGATTGT 1800
AAATGCAACA ATCAGTGCTC TGTGTCTAGC TAAAGGATTG TAAATACACC AATCAGCAC 1859

```

(2) INFORMACIONES PARA LA SEC ID nº: 47:

5 (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 23 pares de bases
 (B) TIPO: Nucleótido
 (C) NÚMERO DE HEBRAS: simple
 (D) CONFIGURACIÓN: Lineal

10

- (ii) TIPO DE MOLÉCULA: ADNc
 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 47:

15 TGATGTGAAC GGCATACTCA CTG

23

(2) INFORMACIONES PARA LA SEC ID nº: 48:

20 (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 24 pares de bases
 (B) TIPO: Nucleótido
 (C) NÚMERO DE HEBRAS: simple
 (D) CONFIGURACIÓN: Lineal

25

- (ii) TIPO DE MOLÉCULA: ADNc
 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 48:

30 CCCAGAGGTT AGGAACTCCC TTTC

24

(2) INFORMACIONES PARA LA SEC ID nº: 49:

35 (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 25 pares de bases
 (B) TIPO: Nucleótido
 (C) NÚMERO DE HEBRAS: simple

	(D) CONFIGURACIÓN: Lineal	
	(ii) TIPO DE MOLÉCULA: ADNc	
	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 49:	
5	GCTAAAGGAG ACTTGTGGTT GTCAG	25
	(2) INFORMACIONES PARA LA SEC ID nº: 50:	
10	(i) CARACTERÍSTICAS DE LA SECUENCIA:	
	(A) LONGITUD: 22 pares de bases	
	(B) TIPO: Nucleótido	
	(C) NÚMERO DE HEBRAS: simple	
15	(D) CONFIGURACIÓN: Lineal	
	(ii) TIPO DE MOLÉCULA: ADNc	
	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 50:	
20	CAACATGGGC ATTCGGATT AG	22

(2) INFORMACIONES PARA LA SEC ID nº: 51:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- 5 (A) LONGITUD: 400 pares de bases
(B) TIPO: Nucleótido
(C) NÚMERO DE HEBRAS: simple
(D) CONFIGURACIÓN: Lineal

- 10 (ii) TIPO DE MOLÉCULA: ADNc
(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 51:

```

GGCTGCTAAA GGAGACTTGT GGTGTCAGA CAATCGCCTA CTTAGGTACC AGGCCTTATT 60
ACTTGAGGGA CTGGTGCTTC AGATGCGCAC TTGTGCAGCT CTTAACCCTA ACTTATGCTG 120
CCCAGAAGGA TCTTTTAGAG GTCCCCTTAG CCAACCCTGA CCTCAACCTA TATATATACT 180
GATGGAAGTT CGTTTGTAGA AAAGGGATTA CAAAGGGNAG GATATNCCAT AGGTTAGTGA 240
TAAAGCAGTA CTTGAAAGTA AGCCTCTTCC CCCCAGGGAC CAGCGCCCCC GTTAGCAGAA 300
CTAGTGGCAC TGACCCCGAG CCTTAGAACT TGGAAAGGGA GGAGGATAAA TGTGTATACA 360
GATAGCAAGT ATGCTTATCT AATCCGAAAT GCCCATGTTG 400
    
```

15 (2) INFORMACIONES PARA LA SEC ID nº: 52:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- 20 (A) LONGITUD: 2389 pares de bases
(B) TIPO: Nucleótido
(C) NÚMERO DE HEBRAS: simple
(D) CONFIGURACIÓN: Lineal

- 25 (ii) TIPO DE MOLÉCULA: ADNc
(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 52:

TCAGGGATAG	CCCCCATCTA	TTTGGTCAGG	CACTGGCCCA	AGATCTAGGG	ACATGCCACT	60
TTTAAGAGCC	ATTTCTCAAG	TCCAGGTACT	CTGGTCCTTC	GGTATGTGGA	TGATTTACTT	120
TTGGCTACCA	GTTCAAGTAGC	CTCATGCCAG	CAGGCTACTC	TAGATCTCTT	GAACCTTCTA	180
GCTAATCAAG	GGTACAGGC	ATCTAGGTTG	AAGGCCAGC	TTTGCCTACA	GCAGGTCAAA	240
TATCTAGGCC	TAATCTTAGC	CAGAGGGACC	AGGGCACTCA	GCAAGGAACA	AATACAGCCT	300
ATACTGGCTT	ATCCTCACCC	TAAGACATTA	AAACAGTTGC	GGGGGTTCCT	TGGAATCACT	360
GGCTTTTGG	TGACTATGGA	TTCCCAGATA	CAGCAAGATT	GGCAGGCCCC	TCTATACTGT	420
AATCAAGGAG	ACTCACGAGG	GCAAGTACTC	ATCTAGTAGA	ATGGGAACCTA	GGGACAGAAA	480
CAGCCTTCAA	AACCTTAAAG	CAGGCCCTAG	TACAATCTCC	AGCTTTAAGC	CTTCCCACAG	540
GACAAAACCTT	CTCTTTATAC	ATCACAGAGA	GGGCAGAGAT	AGCTCTTGGT	GTCCTTATTC	600
AGACTCATGG	GACTACCCCA	CAACCAGTGG	CACACCTAAG	TAAGGAAATT	GATGTAGTAG	660
CAAAAGGCTG	GCCTCACTGT	TTATGGGTAG	CTGTGGTGGT	GGCTGTCTTA	GTGTCAGAAG	720
CTATCAAAAT	AATACAAGGA	AAGGATCTCA	CTGTCTGGAC	TACTCATGAT	GTAATGGCAT	780
ACTAGGTGCC	AAAAGAAGTT	TATGGGTATC	AGACAACCAC	CTGCTTAGAT	ACCAGGGACT	840
ACTCCTGGAG	GATTGGGCTT	CAAGTGCGTT	TTTTGTGGCC	TCAACCCTGC	CACTTTTCTT	900
CCAGAGGATG	GAGAGCCGCT	TGAGCATGCT	TGCCAACAGG	TTGTAGGCCA	GAATTATTCC	960
ACCCGAGATG	ATCTCTTAGA	GTACCCTTAG	CTAATCCTGA	CCTTAACCTA	TATACCAATG	1020
GAAGTTCATT	TGTGGAAAAC	GGGATATGAA	GGGCAGGTTA	TGTCATAGTT	AGTGATGTAA	1080
TCATACTTGC	AAGTAAGCCT	CTTACCCAG	GGGCCAGCAC	TCAGTTAGCA	GAAGTAGTCA	1140
CACCTACCTT	AACCTTAGAA	CTGGGAAAGG	GAAAAAGAA	AAATATGTAT	ACAGATAGTA	1200
AGTATGCTTA	TCTAATCCTA	CATGCCCATG	CTGCAATATG	GAAGGAAAGG	GAGTTCCTAA	1260
CCCTTGGGGG	AACCCCATTT	AAATACCACA	AGGYAAATCA	TGGAGTTATT	GCACGCAGTG	1320
CAAAAACCTCA	AGGAGGTGGC	AGTCTTACAC	TGCCGAAGCY	ATCAAAAAGG	GGAAGGAGAG	1380
GGGAGAACAG	CAGCATTAAGT	GTTTGGCAGA	GGCAGTGAAA	GACCAGCAGA	GAGAAGGAGA	1440
GAGACAACGT	CAACGACAGA	AGGAAGAAG	AGGAGGAGAC	AGAGAGGAAG	AGACAGAGAG	1500
ACAGTTAGTC	CAAGAGAGAG	ACAGAGAGAG	GAAGAGACAG	ACAGAAAGTC	CAAGAGAGAA	1560
GGAAAGAGAG	GAAGAGACCA	AGGAGTCCNA	GAGAGAGAAA	GAGATAGAAG	TAGTAAAGAA	1620
AAAACATTGT	ACCCTATTCC	TTTAAAAGCC	GGGGTATATT	TAAAACCTAT	AATTGATAAT	1680
TGAGTTCTTG	CACCCTCCTC	CAGGGGATYG	CTGGGAGGAA	ACCCTCAACC	GATATGTGAA	1740
AATTGTGGGT	CGTCCCTATG	TCTCAATTAC	CAGCCAATAC	CCCCTTGTTT	TTAGTGTGAA	1800
CGAGGGTGTA	GAGCGCAGAC	AGGGAGACCT	CTGACAATCC	ATACCCTTCC	TATCCAAAAT	1860
CCTTAACCCA	CAGGTTTTC	TAAAAGGGGA	TCTAAATCTT	AATTAATTAC	CATACAAAGG	1920
TCAAACCAGA	TCTAGGAGGA	ACTTCCTTCA	GGACAGGATG	ATAGATGGTT	CCTCCCAGGC	1980
GATTAAAGAA	AATAAAAGA	CACATGGGCA	GCCAGTAAGT	GATAAGGGAA	CACTAGTAGA	2040
AGCAGTTAGG	AGAAGTTGCC	TAATAATTGG	TCTACTCCAA	ATGTGTGAGT	TGTTCCGACT	2100
CAGCCCAAAT	CTTAAAGTAC	TTACAGAATT	AGGGAGGAGC	CATTTACACC	AATTCTAAGT	2160
TAATATGGAC	TGGATGAGGT	TTTATTAATA	GCGAAGGAGA	ATTAAATCCT	AAACTNACAA	2220
GGTTTTC AAC	TAAAGTAAAT	TTTACTAAAA	GCTAACAGTG	TAACATGCAT	TATCCTACTA	2280
CAACACACTC	TCANAGGATT	CCTCAGACAG	TTTACAAGAA	ATAACAAAAT	CTATCTGGTA	2340
AGGATAGTAA	CTACAATCCC	AAATACATTC	TTTGGCAGCA	GTGACTCTC		2389

(2) INFORMACIONES PARA LA SEC ID nº: 53:

- 5 (i) CARACTERÍSTICAS DE LA SECUENCIA:
- (A) LONGITUD: 2448 pares de bases
- (B) TIPO: Nucleótido
- (C) NÚMERO DE HEBRAS: simple
- 10 (D) CONFIGURACIÓN: Lineal
- (ii) TIPO DE MOLÉCULA: ADNc
- (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 53:

TCAGGGATAG	CCCCCATCTA	TTTGATCAGG	CACTAGCCCCA	AGATCTAGGC	CACTTCTGAA	60
GTCCAGGCAT	TCTAGTCCTT	CAGTATGTGG	ATGATTTACT	TTTGGCTACC	AGTTTGGGAAG	120
CCTCATGCCA	GCAGGCTACT	TGAGATCTCT	TGAACTTTCT	AGCTAATCAA	GGGTGTATGG	180
CATCTAAATT	GAAAGTCCAG	CTCTGCCTAC	AACAAGTCAA	ATATCTAGGC	CTAATCTTAG	240
ATAGAAGAAC	CAGGGCCCTC	AGCAAGGAAT	GAATAAAGCC	TATGCTGGCT	TATCGGCACC	300
CTAAGACATT	AAAACAATTG	TGGGGGTTCC	TTGGAATCAC	TGGCTTTTGC	CGACTATGGA	360
TCCCTGGATA	GAGTGAGATA	GCCAGGCCCC	CTCTATTACT	CTTATCAAGG	AGACCCAGAG	420
GGCAAATACT	TATCTAGTAT	TATGGGNACC	AGAGGCAGAA	AAAGCCTTCC	AAACCTTAAA	480
GGAGACCCTA	GTACAAGCTC	CAGCTTTAAG	CCTTCCCACA	GGACAAANCT	TCTCTTTATA	540
TGTCACAGAG	AGAGCAGGAA	TAGCTCCTGG	AGTCCTTACT	CAGACTTTTG	GACGACCCCA	600
CGGCCAGTGG	CRTACCTAAG	TAAGGAAATT	GATGTAGTAG	CAAAAGGCTG	GCCTCACTGT	660
TTATGGGTAG	TTGCGGCTGT	GGCAGTCTTA	CTGTCAAAGG	CTATCAAAAT	AATACAAGGA	720
AAGGATTTCA	CTATCTGGAC	TACTCATGAG	GAAATGGCA	TATTAGGTGC	CAAAGGAAGT	780
TTTTGGCTAT	CAGACAACCA	CCTGCTCAGA	TTCCAGGCAC	TACTGATTGA	GAGACCAGTG	840
CTTTAAATAT	GTATGTGTGT	GTGTGGCCCT	CAACCCTGCC	ACTGTTCTCC	CAGAAGATGG	900
AGAACCAATG	AAGCATTACT	GTCAACAAAT	TAGAGTCCAG	AGTTATGCTG	CCTGAGAGGA	960
TCTCTTAGAA	GTCCCTTAG	CTAATCCTGA	CCTTAACCTA	TATGCTGATG	GAAGTTCACT	1020
TGTGGAGAAT	GGGATACGAA	AAGCACATTA	TGCCATAGTT	AGTGAGGTAA	CAGTACTTGA	1080
AAGTAAGCCT	ATTCCCCCAT	GGACCAGAGC	CCAGTTAGCA	GAAC TAGTGG	CACTTACCCA	1140
AGCCTTAGAA	CTAGGAAAGG	GAAAAATAAT	AAGATGTGTAT	ACAGTAGCA	AGTAGCTTGA	1200
TCTAATCCTA	CATGCCCATG	CTGCAGTATG	GAAAGAAAGG	GAGTTCCTAA	CCTCTGGGGG	1260
AACCCCCATT	AAATACCACA	AGGCAAATCA	TGGAGTTATT	GCATGTAGTG	CAAACCTCA	1320
AGTAGGTGGC	AGTTTTACAC	TGCC TGAAGC	TATGGGGAAG	GAGAGAGGAG	AACAGCAGCA	1380
TAAGTGGCTA	GCAGAGGCAG	CGAAAGACTA	GCAGAGAGGA	AGGGTAGGGG	AAAGACAGAA	1440
AGTCAAAGAA	AAGAAGTCAA	AGACAGACAG	AGAAAGAGAC	AGAGGGAGCC	AGAGAGAAAG	1500
AAAAGAGAGA	ACGAAAGAGA	CAGAATGTCA	AAGAACAGAA	GAGAGAGGCA	GCGCCAGAAG	1560
AGTTAAGAAA	GTGAGAAAGA	GAGATGGAAA	TAGTAAAGAA	AAAACAGTGT	ACCCTATTCC	1620
TTTAAAGGCC	AGGGTAAATT	TAAAACGTAT	AATTTTATAA	TTGGAAGGTC	TTCTCCATAA	1680
CCCTATAACA	TAAAATACC	ACCTTGTGTG	CAGTGTA AAC	AAGAGCATAG	CCCAAAAGCA	1740
CTGAGGCCAC	TGACAACCCA	TAGCCTTCCT	ATCAAAAATC	CTTAACCTCTG	CAGGTTTCCT	1800
AACAGGGGAT	CTAAATCTCA	ACTAATCACC	ATACAATGGT	CCGACCAGAC	CTAGGAGCGA	1860
CTCCCCCTCAG	GACAGAAGGA	TGGATGGTTC	CTCCCAGGCC	ATTAAGGGAA	AGAGACACAA	1920
TGGGTATTCA	GTAAGTGATA	AGGGA ACTCT	TGTAGAAGCA	GTTAGGAAGA	TTGCCTAATA	1980
TTTGGTCTGC	TCAAATGTGC	CAGCTGTTTG	CACTCAGCTA	AACCTTAAAT	TACTTACAGA	2040
ATTAGGAAGG	AGCCATCTAT	ACCAATTCTG	AGTTAATATG	AGCTGAACAA	GTTCTTATTA	2100
ATAGCAAAGA	ATCATTGAAA	TCTCAAACCT	GCAAAGTTTT	CAACAAAAGT	AAAGTTTGCT	2160
GAAAGTTAGC	AGTGTAAACAT	GTATTATCCT	AACTTCTAAT	CTTGTGGAAA	TCAGACCCTA	2220
TCAGTGCCCC	TCAAAGCTGA	AGTCCATCAG	CATATGGCCA	TACA ACTAAT	ACCCCTATTT	2280
ATAGGGTTAG	GAATGGCCAC	TGCTACAGGA	ATGGGAGTAA	CAGGTTTATC	TACTTCATTA	2340
TCCTATTACC	ACACACTCTT	AAAGGATTTC	TCAGACAGTT	TACAAGAAAT	AACAAAATCT	2400
ATCCTTACTC	TNTARTCCCA	AATAGRTTCT	TTGGCAGCAG	TGACTCTC		2448

(2) INFORMACIONES PARA LA SEC ID nº: 54:

- 5 (i) CARACTERÍSTICAS DE LA SECUENCIA:
- (A) LONGITUD: 21 pares de bases
- (B) TIPO: Nucleótido
- (C) NÚMERO DE HEBRAS: simple
- 10 (D) CONFIGURACIÓN: Lineal
- (ii) TIPO DE MOLÉCULA: ADNc
- (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 54:

15 CCTGAGTTCT TGCACTAACC C

21

(2) INFORMACIONES PARA LA SEC ID nº: 55:

- 20 (i) CARACTERÍSTICAS DE LA SECUENCIA:
- (A) LONGITUD: 23 pares de bases
- (B) TIPO: Nucleótido
- (C) NÚMERO DE HEBRAS: simple
- (D) CONFIGURACIÓN: Lineal
- 25 (ii) TIPO DE MOLÉCULA: ADNc
- (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 55:

GTCCGTTGGG TTTCCTTACT CCT23

(2) INFORMACIONES PARA LA SEC ID nº: 56:

5 (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 1196 pares de bases

(B) TIPO: Nucleótido

(C) NÚMERO DE HEBRAS: simple

10 (D) CONFIGURACIÓN: Lineal

(ii) TIPO DE MOLÉCULA: ADNc

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 56:

TTCCTGAGTT	CTTGCACTAA	CCTCAAATGA	GAGAAGTGCC	GCCATAACTG	CAACCCAAGA	60
TTTTGGCGAT	CCCTGGTATC	TCAGTCAGGT	CAATGACAGG	ATGACACACAG	AGGAAAGATA	120
ATGATTCCCC	ACAGGCCAGC	AGGCAGTTCC	CAGTGTAGAC	CCTCATTAGG	ACACAGAATC	180
AGAACATGGA	GATTGGTGCC	GCAGACATTT	GCTAACTTGC	GTGCTAGAAG	GAATAAGGAA	240
AAC TAGGAAG	ATATGAATTA	TTCAATGATG	TCCACTATAA	CACAGGGGAA	AGGAAGAAAA	300
TCCTACTGCC	TTTCTGGAGA	GACTAAGGGA	GGCATTGAGG	AAGCATACCA	GGCAAGTGGA	360
CATTGGAGGC	TCTGGAAGAG	GGAAAAGTTG	GGAAAAGTAT	ATGTCTAATA	GGGCTTGCTT	420
CCAGTGTGGT	CTACAAGGAC	ACTTTAAAAA	AGATTGTCCA	ATAGAAATAA	GCCACCACCT	480
CGTCCATGCC	CCTTATGTCA	AGGGAATCAC	TGGAAGGCC	ACTGCCCCAG	GGGATGAAGG	540
TCCTCTGAGT	CAGAAGCCAC	TAACCAGATG	ATCCAGCAGC	AGGACTGAGG	GTGCCCCGGG	600
CAAGCGCCAG	CCCATGCCAT	CACCCTCACA	GAGCCCCAGG	TATGCTTGAC	CATTGAGGGT	660
CAGAAGGGTA	CTGTCTCCTG	GACACTGGCG	GGCCTTCTCA	GTCTTACTTT	CCTGTCCTGG	720
ACAAC TGTC	TCCAGATCTG	TCACTGTCCG	AGGGGTCCTA	GGACAGCCAG	TCACTAGATA	780
CTTCTCCCG	CCACTAAGTT	GTGACTGGGG	AACTTTACTC	TTCCACATGC	TTTTCTAATT	840
ATGCCTGAAA	GCCCCACTCT	CTTGTTAGGG	GAGAGACATT	CTAGCAAAAG	CAGGGGCCAT	900
TATACATGTG	AATATAGGAG	AAGGAACAAC	TGTTTGTGTG	CCCCTGCTTG	AGGAAGGAAT	960
TAATCCTGAA	GTCCGGGCAA	CAGAAGGACA	ATATGGACAA	GCAAAGAATG	CCCGTCCTGT	1020
TCAAGTTAAA	CTAAAGGATT	CCACCTCCTT	TCCCTACCAA	AGGCAGTACC	CCCTCAGACC	1080
CGAGACCCAA	CAAGAACTCC	AAAAGATTGT	AAAGGACCTA	AAAGCCCAAG	GCCTAGTAAA	1140
ACCAAGCAAT	AGCCCTTGCA	AGACTCCAAT	TTTAGGAGTA	AGGAAACCCA	ACGGAC	1196

15

(2) INFORMACIONES PARA LA SEC ID nº: 57:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 2391 pares de bases

(B) TIPO: Nucleótido

(C) NÚMERO DE HEBRAS: simple

(D) CONFIGURACIÓN: Lineal

25

(ii) TIPO DE MOLÉCULA: ADNc

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 57:

ATGATCCAGC	AGCAGGACNG	AGGGTGCCCG	GGGCAAGCGC	CAGCCCATGC	CATCACCCCTC	60
ACAGAGCCCC	AGGTATGCTT	GACCATTGAG	GGTCAGAAGG	GTNACTGTCT	CCTGGACACT	120
GGCGGNGCCT	TCTCAGTCTT	ACTTTCCTGT	CCTGGACAAC	TGTCCTCCAG	ATCTGTCACT	180
GTCCGAGGGG	TCCTAGGACA	GCCAGTCACT	AGATACTTCT	CCCAGCCACT	AAGTTGTGAC	240
TGGGGAACCT	TACTCTTCCC	ACATGCTTTT	CTAATTATGC	CTGAAAGCCC	CACTCTCTTG	300
TTGGGGAGAG	ACATTCTAGC	AAAAGCAGGG	GCCATTATAC	ATGTGAATAT	AGGAGAAGGA	360
ACAACGTGTT	GTTGTCCCCT	GCTTGAGGAA	GGAATTAATC	CTGAAGTCCG	GGCAACAGAA	420
GGACAATATG	GACAAGCAAA	GAATGCCCGT	CCTGTTCAAG	TTAAACTAAA	GGATTCCACC	480
TCCTTTCCCT	ACCAAAGGCA	GTACCCCTC	AGACCCGAGA	CCCAACAAGA	ACTCCAAAAG	540
ATTGTAAAGG	ACCTAAAAGC	CCAAGGCCTA	GTAAAACCAA	GCAATAGCCC	TTGCAAGACT	600
CCAATTTTAG	GAGTAAGGAA	ACCCAACGGA	CAGTGGAGGT	TAGTGCAAGA	ACTCAGGATT	660
ATCAATGAGG	CTGTTGTTCC	TCTATACCCA	GCTGTACCTA	ACCCTTATAC	AGTGCTTTCC	720
CAAATACCAG	AGGAAGCAGA	GTGGTTTACA	GTCTGGACC	TTAAGGATGC	CTTTTTCTGC	780
ATCCCTGTAC	GTCCTGACTC	TCAATTCTTG	TTTGCCTTTG	AAGATCCTTT	GAACCCAACG	840
TCTCAACTCA	CCTGGACTGT	TTTACCCCAA	GGGTTCAAGG	ATAGCCCCCA	TCTATTTGGC	900
CAGGCATTAG	CCCAAGACTT	GAGTCAATTC	TCATACCTGG	ACACTCTTGT	CCTTCAGTAC	960
ATGGATGATT	TACTTTTAGT	CGCCCGTTCA	GAAACCTTGT	GCCATCAAGC	CACCCAAGAA	1020
CTCTTAACCT	TCCTCACTAC	CTGTGGCTAC	AAGGTTTCCA	AACCAAAGGC	TCGGCTCTGC	1080
TCACAGGAGA	TTAGATACTN	AGGGCTAAAA	TTATCCAAAG	GCACCAGGGC	CCTCAGTGAG	1140
GAACGTATCC	AGCCTATACT	GGCTTATCCT	CATCCCAAAA	CCCTAAAGCA	ACTAAGAGGG	1200
TTCCTTGGCA	TAACAGGTTT	CTGCCGAAAA	CAGATTCCCA	GGTACASCCC	AATAGCCAGA	1260
CCATTATATA	CACTAATTAN	GGAACTCAG	AAAGCCAATA	CCTATTTAGT	AAGATGGACA	1320
CCTACAGAAG	TGGCTTTCCA	GGCCCTAAAG	AAGGCCCTAA	CCCAAGCCCC	AGTGTTTCAGC	1380
TTGCCAACAG	GGCAAGATTT	TTCTTTATAT	GCCACAGAAA	AAACAGGAAT	AGCTCTAGGA	1440
GTCCTTACCG	AGGTCTCAGG	GATGAGCTTG	CAACCCGTGG	TATACCTGAG	TAAGGAAATT	1500
GATGTAGTGG	CAAAGGGTTG	GCCTCATNGT	TTATGGGTAA	TGGNGGCAGT	AGCAGTCTNA	1560
GTATCTGAAG	CAGTTAAAT	AATACAGGGA	AGAGATCTTN	CTGTGTGGAC	ATCTCATGAT	1620
GTGAACGGCA	TACTCACTGC	TAAAGGAGAC	TTGTGGTTGT	CAGACAACCA	TTTACTTAA	1680
TATCAGGCTC	TATTACTTGA	AGAGCCAGTG	CTGNAGCTGC	GCACTTGTGC	AACTCTTAA	1740
CCCAAACCTA	TGCTGCCCCAG	AAGGATCTTT	NTAGAGGTCC	CCTTAGCCAA	CCCTGACCTC	1800
AACTATATAT	ATACTGATGG	AAGTTCGTTT	GTAGAAAAGG	GATTACAAAG	GGNAGGATAT	1860
NCCATAGGTG	TTAGTGATAA	AGCAGTACTT	GAAAGTAAGC	CTCTTCCCCC	CCAGGGACCA	1920
GCGCCCCCGT	TAGCAGAACT	AGTGGCACTG	ACCCCGCGAG	CCTTAGAACT	TTGGAAAGGG	1980
AGGAGATATA	ATGTGTATAC	AGATAGCAAG	TATGCTTATC	TAATCCGAAA	TGCCCATGTT	2040
GTTTATCTAA	TCCGAAATGC	CCATGTTGCA	ATATGGAAAG	AAAGGGAGTT	CCTAACCTCT	2100
GGGGGAACCC	CCATTAAATA	CCACAAGTTA	ATCATGGAGT	TATTGCACAC	AGTGCAAAAA	2160
CTCAAGGAGG	TGGAAGTCTT	ACACTGCCAA	AGCCATCAGA	AAAGGGAAAG	GGGAGAAGAG	2220
CAGCATAAGT	GGCTACAGAG	GCAAGGAAAG	ACTAGCAGAA	AGGAAAGAGA	GAAAGAGACA	2280
GAAAGTCAGA	GAGAGAGAGA	GGAAGAGACA	GAGCACAAG	AGGGAGTCAG	AGAGAGAGAG	2340
AGACAGAGAG	TCAGAGAGAA	GGAAAGAGAG	AGAGGAAGAG	ACAAAGAATG	A	2391

(2) INFORMACIONES PARA LA SEC ID nº: 58:

5 (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 1722 pares de bases
- (B) TIPO: Nucleótido
- (C) NÚMERO DE HEBRAS: simple
- (D) CONFIGURACIÓN: Lineal

10

- (ii) TIPO DE MOLÉCULA: ADNc
- (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 58:

TGGAGAATAG	CAGCATAAGT	TGGCTGGCAG	AAGTAGGGAA	AGACAGCAAG	AAGTAAAGAA	60
AAAAARGAGA	AAGTCAGAGA	AAGAAAAAAA	GAGAGGAAGA	AACAAAGAAG	AACCTGAAGA	120
GAGAAAGAAG	TAGTAAAGAA	AAAACAGTAT	ACCCTATTCC	TTTAAAAGCC	AGGGTAAATT	180
TCTGTCTACC	TAGCCAAGGC	ATATTCTTCT	TATGTGGAAC	ATCAACCTAT	ATCTGCCCTCC	240
CCACTAAGTG	GACAGGCACC	TGAACCTTAG	TCTTTCTAAG	TCCCAACATT	AACATTGCCC	300
CAGGAAATCA	GACCTTATTG	GTACCTGTCA	AAGCTAAAGT	CCCGTCAGTG	CAGAGCCATA	360

15

CAACTAATAT	CCCTATTAT	AGGGTTAGGA	ATGGCTACTG	CTACAGGAAC	TGGAATAGCC	420
GGTTTATCTA	CTTCATTATC	CTACTACCAT	ACACTCTCAA	AGAATTTCTC	AGACAGTTTG	480
CAAGAAATAA	TGAAATCTAT	TCTTACTTTA	CAATCCCAAT	TAGACTCTTT	GGCAGCAATG	540
ACTCTCCAAA	ACCGCCGAGG	CCCACACCTC	CTCACTGCTG	AGAAAGGAGG	ACTCTGCACC	600
TTCTTAGGGG	AAGAGTGTG	TTTTTACACT	AACCAGTCAG	GGATAGTACG	AGATGCCACC	660
TGGCATTTAC	AGGAAAGGGC	TTCTGATATC	AGACAATGCC	TTTCAAATC	TTATACCAAC	720
CTCTGGAGTT	GGGCAACATG	GCTTCTTCCA	TTTCTAGGTC	CCATGGCAGC	CATCTTGCTG	780
TTACTCACCT	TTGGGCCCTG	TATTTTTAAG	CTTCTTGTC	AATTTGTTTC	CTCTAGGATC	840
GAAGCCATCA	AGCTACAGAT	GGTCTTACAA	ATGGAACCCC	AAATGAGTTC	AACTAACAA	900
TTCTACCAAG	GACCCCTGGA	ACGATCCACT	GGCACTTCCA	CTAGCCCTAG	GATTCCCTC	960
TGGAAGACAC	TACAACTGCA	GGGCCCCCTC	TTTGCCCCCTA	TCCAGCAGGA	AGTAGCTAGA	1020
GCGGTCATCG	GCCAAATTCC	CAACAGCAGT	TGGGGTGTCC	TGTTTAGAGG	GGGGATTGAA	1080
GAGGTGACAG	CCTGCTGGCA	GCCTCACAGC	CCTCGTTGGY	TCTCAGTGCC	TCCTCAGCCT	1140
TGGTGCCAC	TCTGGCCGTG	CTTGAGGAGC	CCTTCAGCCT	GCCACTGCAC	TGTGGGAGCC	1200
TCTTTCTGGG	CTGGACAAGG	CCGGAGCCAG	CTCCCTCAGC	TTGCAGGGAG	GTATGGAGGG	1260
AGAGATGCAG	GCGGAACCA	GGGCTGCGCA	TGGCGCTTGC	GGGCCAGCAT	GAGTTCCAGG	1320
TGGGCGTGGG	CTCGGCGGGC	CCCACACTCG	GGCAGTGAGG	GGCTTAGCAC	CTGGGCCAGA	1380
CAGATGCTGT	GCTCAACTTC	TTGCTGGGCT	CTTAGCTGCC	TTCCCGTGG	GGCAGGGCTY	1440
CGGGAACMTG	CAGCCTGCCC	ATGCTTGAGC	CCCCACCCC	GCCGTGGGTT	CYTGCACAGC	1500
CCAAGCTTCC	CGGACAAGCA	CCACCCCTTA	TCCACGGTGC	CCAGTCCCAT	CAACCACCCA	1560
AGGGTTGAGG	AGTGGGGGCA	CACAGCGGAG	GATTGGCAGG	CAGTTCCACT	TGCGGCCCTG	1620
GTGCGGGATC	CACCTCGTGA	AGCCAGCTGG	GCTCCTGAGT	CTGCTGGGGA	CTTGAGAAAT	1680
CTTTATGTCT	AGCTAAGGGA	TTGTAAATAC	ACCAATCAGC	AC		1722

(2) INFORMACIONES PARA LA SEC ID nº: 59:

5 (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 495 pares de bases
 (B) TIPO: Nucleótido
 (C) NÚMERO DE HEBRAS: simple
 (D) CONFIGURACIÓN: Lineal

10

- (ii) TIPO DE MOLÉCULA: ADNc
 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 59:

CTTCCCCAAC	TAATAAGGAC	CCCCCTTTCA	ACCCAAACAG	TCCAAAAGGA	CATAGACAAA	60
GGAGTAAACA	ATGAACCAA	GAGTGCCAAT	ATTCCCTGGT	TATGCACCCT	CCAAGCGGTG	120
GGAGAAGAAT	TCGGCCGAGC	CAGAGTGAT	GTACCTTTTT	CTCTCTCACA	CTTGAAGCAA	180
ATTAAATAG	ACNTAGGTNA	ATTNTCAGAT	AGCCCTGATG	GYTATATTGA	TGTTTTACAA	240
GGATTAGGAC	AATCCTTTGA	TCTGACATGG	AGAGATATAA	TATTACTGCT	AAATCAGACG	300
CTAACCTCAA	ATGAGAGAAG	TGCTGCCATA	ACTGGAGCCC	GAGAGTTTGG	CAATCTCTGG	360
TATCTCAGTC	AGGTCAATGA	TAGGATGACA	ACGGAGGAAA	GAGAACGATT	CCCCACAGGG	420
CAGCAGGCAG	TTCCAGTGT	AGCTCCTCAT	TGGGACACAG	AATCAGAACA	TGGAGATTGG	480
TGCCGACAGC	ATTTA					495

15

(2) INFORMACIONES PARA LA SEC ID nº: 60:

20 (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 2503 pares de bases
 (B) TIPO: Nucleótido
 (C) NÚMERO DE HEBRAS: simple
 (D) CONFIGURACIÓN: Lineal

25

- (ii) TIPO DE MOLÉCULA: ADNc
 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 60:

CCAAGAACC	ACCAATTCCG	GANCACATT	TGGCGACCAC	GAAGGGACTT	TCGCATATCG	60
CCAAGCGGTG	AGACAATAGC	CGAGCGGTGA	GACCTTTCCC	AATCGCCAAG	CAGTGAGTAC	120
CATCAGACCC	CTTTCACCTG	CTATTCTGTC	CTATCTTTCT	TTAGAATTTCG	GGGGCTAAAT	180
ACCGGGCATC	TGTCAGCCAT	TTAAAAGTGA	CTAGCGGGCC	GCCGGACTAA	AGACACGGGT	240

GTCAAGCTTT	CTGGGAAAGG	GCTCTCTAAC	AACCCCCAAC	TCTTTGGAGT	TGGGACCGTT	300
GGTTTGCCTA	GAACCAGCTT	CCGCTTTTCC	TGTACTTCTG	GGCTGAGCCG	TGGGTTGACA	360
GTGAAGGAAA	GCCATGCATC	TCCGGGGTCT	CGMCAACATG	TTGGTTGACC	CTGCGGCCAT	420
GAGTGGAAC	CTCAAAAGCA	TGTCGCCCAA	GCGACACTCG	CCTATCTATC	CTATCTATCC	480
TGACCCCTGC	CCTCTGGGTC	CTAATGCCTG	CCAGACAAAC	TTCTCTCGC	CTCTCTTCTC	540
TGAAGCTAGA	ACCGCTTCTA	AAAATGCTA	CCTGGTCTCT	GGTGCTTTTC	CTARTTCTC	600
CTATAAAGAA	TGAWTTCTAG	TATTAACTC	CAGGACTCTG	TTACCTTCTT	TAGGCACCCG	660
GGCTCACCAA	TCAGAAAGAC	ACAGTTTTTG	CCCAGGCC	CATCGTAGTG	GGGACTACCT	720
GGATTTTATG	GATCCCTCCT	CAGACTAACA	GGCCTAACAA	AAGTTATTCC	TGAAGCTAGG	780
ATATGGGGAG	CCTCAGAAAT	TGTATCCCTC	CTATTCATAT	AAGTGAGAAC	AAAAGGTGTC	840
ACTCTTCCAA	CCCTGAAGAT	CCCCTCCCTC	CCTCAGGGTA	TGGCCCTCCA	TTTCATTTTTT	900
GTGGCATAAC	ATCTTTATAG	GATGGGGTAA	AGTCCAATA	CTAACAGGAG	AATGCTTAGG	960
ACTCTAACAG	GTTTTTGAGA	ATGCGTCAGT	AAGGGCCACT	AAATCTGATT	TTTCTCAGTC	1020
GGTCTCTCTT	TGGGTCTAGG	AGGACAGGCA	AGGTGTGCA	GGTTTTCGAG	AATGCGTCAG	1080
TAAGGACCAC	TAAATCCGAC	CTTCTCGGT	CCTCCATGTG	GTCTGGGAGG	AAAAC TAGTG	1140
TTTCTGCTGC	TGCGTCCGTG	AGCGCAACTA	TTCAAGTCAG	CAGGGTCCAG	GGACCGTTGC	1200
AGGTTCTTGG	GCAGGGGTG	TTTCTGCTGC	TGCATTGGTG	AATGCAACTA	TTCTGATCAG	1260
CAGGGTCCCA	GGACCATTGC	AGGTCCTTGG	GCAGGGAGAG	AAACAAAACA	AACCAAAACT	1320
GTGGGCGGTT	TTGTCTTTCA	TATGGGAAAC	ACTCAGGCAT	CAACAGGTTT	ACCCTTGAAA	1380
TGCATCCTAA	GCCATTGGGA	CCAATTGAC	CCACAAACCC	TGAAAAAGAG	GAGGCTCATT	1440
TTTTCTTGCA	CTACGGCTTG	GCCCCAATAT	TCTCTTGTG	ATGGGGAAAA	ATGGCCACCT	1500
GAGGGAAGCA	CAAATTACAA	TAYTATCCTA	CAGCYTGATC	TTTCTGTAA	GAGGGAAGGC	1560
AAATGGAGTG	AATACCTTAT	GTCCAAAGCTT	TCTTTTCATT	GAGGGAGAAT	ACACAACTAT	1620
GCAAAGCTTG	CAATTTACAT	CCCACAGGAG	GACCTTTCAG	CTTACCCCA	TATCCTAGCC	1680
TCCCTATAGC	TTCCCTTCCT	ATTGATGATA	CTCCTCCTCT	AATCTCCCCT	GCCCAGAAGG	1740
AAATAAGCAA	AGAAATCTCC	AAAGGTCCAC	AAAAACCCCC	GGGCTATCGG	TTATGTCCCT	1800
TCAAGYTGTA	GGGGGAGGGG	AATTTGGCCC	AACCCGGGTG	CATGTCCCTT	CTCCCTCTCT	1860
GATTTAAAGC	AGATCAAGGC	AGACCTGGGG	AAGTTTTCAG	ATGATCCTGA	TAGGTACATA	1920
GATGTCTTAC	AGGGTCTAGG	GCAAACCTTT	GACCTCACTT	GGAGAGACGT	CATGCTACTG	1980
TTAGATCAAA	CCCTGGCCTT	TAATGAAAAG	AATGCGGCTT	TAGCTGCAGC	CTGAGAGTTT	2040
GGAGATACCT	GGTATCCTAG	TCAAGTAAAT	GAAAGAATGA	CAGCCGAAGA	AAGGGACAAC	2100
TTCTTTACTG	GTCAGCAACC	CATCCCAGT	ATGGATCCCC	ACTGGGACTT	TGACTCAGAT	2160
CATGGGGACT	GGAGTCGTAA	ACATCTGTTG	ATCTGTGTTT	TGGAAGGACT	AAGGAGAATT	2220
GGGAAAAAGC	CCATGAATTA	TTCAATGATA	TCCACCATAA	CCCAGGGAAG	GGAAGAAAAT	2280
CCTTCTGCCT	TCCTCGAGCG	GCTACAAGAG	GCCTTAAGAA	AATATACTCC	CCTGTACCCC	2340
GAATCACTCG	AGGGTCAATT	GATTCTAAAA	GATAAGTTTA	TTACCCAATC	AGCCACAGAT	2400
ATCAGGAGAA	AGCTCCAAAA	GCAAGCCCTG	AGCCTGAACA	AAATCTAGAG	ACATTATTAA	2460
ACCTGGCAAC	CTTGGTGTTT	TATAATAGGG	ACCAAGAGGA	ACA		2503

(2) INFORMACIONES PARA LA SEC ID nº: 61:

- 5 (i) CARACTERÍSTICAS DE LA SECUENCIA:
- (A) LONGITUD: 1167 pares de bases
- (B) TIPO: Nucleótido
- (C) NÚMERO DE HEBRAS: simple
- 10 (D) CONFIGURACIÓN: Lineal
- (ii) TIPO DE MOLÉCULA: ADNc
- (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 61:

AAGGAACTC AGAAAGCCAA TACCCATTTA GTAAGATGGA CACCAGAAGC AGAAGCAGCT 60
TTCCAGGCC TAAAGAAATC CCTAACCCAA GCCCCAGTGT TAAGCTTGCC AACGGGGCAA 120
GACTTTTCTT TATATGTCAC AGAAAAACAG GAATAGCTCT AGGAGTCCTT ACACAGGTCC 180
AAGGGACAAG CTTGCAACCT GTGGCATAACC TGAGTAAGGA AACTGATGTA NTGGCAAAGG 240
GTTGGCCTCA TTGTTTACAG GTAGGGCAGC AGTAGCAGTC TTAGTTTCTG AAACAGTTAA 300
AATAATACAG GGAAGAGATC TTACTGTGTG GACATCTCAT GATGTGAACG GCATACTCAC 360
TGCTAAAGAG GACTTGTGGC TGTCAGACAA CCATTTACTT AAATAGCAGG TTCTATTACT 420
TGAAGTGCCA GTGCTGCGAC TGCACATTTG TGCAACTCTT AACCAGCCA CATTCTTCC 480
AGACAATGAA GAAAAGATAG AACATAACTG TCAACAAGTA ATTGCTCAAA CCTATGCTGC 540
TCGAGGGGAC CTTCTAGAGG TTCCCTTGAC TGATCCCGAC CTCAACTTGT ATACTGATGG 600
AAGTTCCCTG GCAGAAAAAG GACTTTGAAA AGCGGGGTAT GCAGTGATCA GTGATAATGG 660

AATACTTGAA AGTAATCGCC TCACTCCAGG AACTAGTGCT CACCTGGCAG AACTAATAGC 720
CCTCACTTGG GCACTAGAAT TAGGAGAAGG AAAAAGGGTA AATATATATT CAGACTCTAA 780
GTATGCTTAC CTAGTCCTCC ATGCCCATGC AGCAATATGG AGAGAGAGGG AATTCCTAAC 840
TTCTGAGGGA ACACCTATCA ACCATCAGGG AGCCATTAG GAGATTATTA TTGGCTGTAC 900
AGAAACCTAA AGAGGTGGCA GTCTTACACT GCCAGGGTCA TCAGGAAGAA GAGGAAAGGG 960
AAATAGAAGG CAATCGCCAA GCGGATATTG AAGCAAAAAA AGCCGCAAGG CAGGACTCTC 1020
CATTAGAAAT GCTTATAGAA GGACCCCTAG TATGGGGTAA TCCCCTCTGG GAAACCAAGC 1080
CCCAGTACTC AGCAGGAAAA ATAGAATAGG AAACCTCACA AGGACATACT TTCCTCCCCT 1140
CCAGATGGCT AGCCACTGAG GAAGGAA 1167

(2) INFORMACIONES PARA LA SEC ID nº: 62:

5 (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 78 pares de bases
 (B) TIPO: Nucleótido
 (C) NÚMERO DE HEBRAS: simple
 (D) CONFIGURACIÓN: Lineal

10

- (ii) TIPO DE MOLÉCULA: ADNc
 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 62:

TCCAAAGGCA CCAGGGCCCT CAGTGAGGAA CGTATCCAGC CTATACTGGC TTATCCTCAT 60
CCCCAAACCC TAAAGCAA 78

15

(2) INFORMACIONES PARA LA SEC ID nº: 63:

20 (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 26 restos de aminoácidos
 (B) TIPO: aminoácidos

20

- (ii) TIPO DE MOLÉCULA: péptido
 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 63:

25

Ser Lys Gly Thr Arg Ala Leu Ser Glu Glu Arg Ile Gln Pro Ile Leu
1 5 10 15
Ala Tyr Pro His Pro Lys Thr Leu Lys Gln
20 25

(2) INFORMACIONES PARA LA SEC ID nº: 64:

30

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 28 pares de bases
 (B) TIPO: Nucleótido
 (C) NÚMERO DE HEBRAS: simple
 (D) CONFIGURACIÓN: Lineal

35

- (ii) TIPO DE MOLÉCULA: ADNc

	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 64:	
	AAATGTCTGC GGCACCAATC TCCATGTT	28
5	(2) INFORMACIONES PARA LA SEC ID nº: 65:	
	(i) CARACTERÍSTICAS DE LA SECUENCIA:	
10	(A) LONGITUD: 30 pares de bases	
	(B) TIPO: Nucleótido	
	(C) NÚMERO DE HEBRAS: simple	
	(D) CONFIGURACIÓN: Lineal	
	(ii) TIPO DE MOLÉCULA: ADNc	
15	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 65:	
	AAGGGGCATG GACGAGGTGG TGGCTTATTT	30
	(2) INFORMACIONES PARA LA SEC ID nº: 66:	
20	(i) CARACTERÍSTICAS DE LA SECUENCIA:	
	(A) LONGITUD: 21 pares de bases	
25	(B) TIPO: Nucleótido	
	(C) NÚMERO DE HEBRAS: simple	
	(D) CONFIGURACIÓN: Lineal	
	(ii) TIPO DE MOLÉCULA: ADNc	
30	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 66:	
	GGAGAAGAGC AGCATAAGTG G	21
	(2) INFORMACIONES PARA LA SEC ID nº: 67:	
35	(i) CARACTERÍSTICAS DE LA SECUENCIA:	
	(A) LONGITUD: 25 pares de bases	
	(B) TIPO: Nucleótido	
40	(C) NÚMERO DE HEBRAS: simple	
	(D) CONFIGURACIÓN: Lineal	
	(ii) TIPO DE MOLÉCULA: ADNc	
	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 67:	
45	GTGCTGATTG GTGTATTTAC AATCC	25
	(2) INFORMACIONES PARA LA SEC ID nº: 68:	
50	(i) CARACTERÍSTICAS DE LA SECUENCIA:	
	(A) LONGITUD: 34 pares de bases	
	(B) TIPO: Nucleótido	
	(C) NÚMERO DE HEBRAS: simple	
55	(D) CONFIGURACIÓN: Lineal	
	(ii) TIPO DE MOLÉCULA: ADNc	
	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 68:	
60	GACTCGCTGC AGATCGATTT TTTT TTTT	34
	(2) INFORMACIONES PARA LA SEC ID nº: 69:	
	(i) CARACTERÍSTICAS DE LA SECUENCIA:	
65	(A) LONGITUD: 30 pares de bases	
	(B) TIPO: Nucleótido	

	(C) NÚMERO DE HEBRAS: simple	
	(D) CONFIGURACIÓN: Lineal	
5	(ii) TIPO DE MOLÉCULA: ADNc	
	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 69:	
	GCCATCAAGC CACCCAAGAA CTCTTAAGTT	30
10	(2) INFORMACIONES PARA LA SEC ID nº: 70:	
	(i) CARACTERÍSTICAS DE LA SECUENCIA:	
	(A) LONGITUD: 30 pares de bases	
	(B) TIPO: Nucleótido	
15	(C) NÚMERO DE HEBRAS: simple	
	(D) CONFIGURACIÓN: Lineal	
	(ii) TIPO DE MOLÉCULA: ADNc	
20	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 70:	
	CCAATAGCCA GACCATTATA TACTACTAATT	30
	(2) INFORMACIONES PARA LA SEC ID nº: 71:	
25	(i) CARACTERÍSTICAS DE LA SECUENCIA:	
	(A) LONGITUD: 23 pares de bases	
	(B) TIPO: Nucleótido	
	(C) NÚMERO DE HEBRAS: simple	
30	(D) CONFIGURACIÓN: Lineal	
	(ii) TIPO DE MOLÉCULA: ADNc	
	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 71:	
35	GCCATAACTG CAACCCAAGA GTT	23
	(2) INFORMACIONES PARA LA SEC ID nº: 72:	
40	(i) CARACTERÍSTICAS DE LA SECUENCIA:	
	(A) LONGITUD: 23 pares de bases	
	(B) TIPO: Nucleótido	
	(C) NÚMERO DE HEBRAS: simple	
45	(D) CONFIGURACIÓN: Lineal	
	(ii) TIPO DE MOLÉCULA: ADNc	
	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 72:	
50	GGACGAGGTG GTGGCTTATT TCT	23
	(2) INFORMACIONES PARA LA SEC ID nº: 73:	
	(i) CARACTERÍSTICAS DE LA SECUENCIA:	
55	(A) LONGITUD: 25 pares de bases	
	(B) TIPO: Nucleótido	
	(C) NÚMERO DE HEBRAS: simple	
	(D) CONFIGURACIÓN: Lineal	
60	(ii) TIPO DE MOLÉCULA: ADNc	
	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 73:	
	AACTTGCGTG CTAGAAGGAC TAAGG	25
65	(2) INFORMACIONES PARA LA SEC ID nº: 74:	

	(i) CARACTERÍSTICAS DE LA SECUENCIA:	
	(A) LONGITUD: 24 pares de bases	
	(B) TIPO: Nucleótido	
5	(C) NÚMERO DE HEBRAS: simple	
	(D) CONFIGURACIÓN: Lineal	
	(ii) TIPO DE MOLÉCULA: ADNc	
10	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 74:	
	AACTTTTCCC TTTTCCAGAT CCTC	24
	(2) INFORMACIONES PARA LA SEC ID nº: 75:	
15	(i) CARACTERÍSTICAS DE LA SECUENCIA:	
	(A) LONGITUD: 22 pares de bases	
	(B) TIPO: Nucleótido	
20	(C) NÚMERO DE HEBRAS: simple	
	(D) CONFIGURACIÓN: Lineal	
	(ii) TIPO DE MOLÉCULA: ADNc	
	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 75:	
25	GCATACCAGG CAAGTGGACA TT	22
	(2) INFORMACIONES PARA LA SEC ID nº: 76:	
30	(i) CARACTERÍSTICAS DE LA SECUENCIA:	
	(A) LONGITUD: 25 pares de bases	
	(B) TIPO: Nucleótido	
	(C) NÚMERO DE HEBRAS: simple	
35	(D) CONFIGURACIÓN: Lineal	
	(ii) TIPO DE MOLÉCULA: ADNc	
	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 76:	
40	CTGTCCGTTG GGTTTCCTTA CTCCT	25
	(2) INFORMACIONES PARA LA SEC ID nº: 77:	
	(i) CARACTERÍSTICAS DE LA SECUENCIA:	
45	(A) LONGITUD: 24 pares de bases	
	(B) TIPO: Nucleótido	
	(C) NÚMERO DE HEBRAS: simple	
	(D) CONFIGURACIÓN: Lineal	
50	(ii) TIPO DE MOLÉCULA: ADNc	
	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 77:	
	GAGGCTCTGG AAAAGGGAAA AGTT	24
55	(2) INFORMACIONES PARA LA SEC ID nº: 78:	
	(i) CARACTERÍSTICAS DE LA SECUENCIA:	
60	(A) LONGITUD: 25 pares de bases	
	(B) TIPO: Nucleótido	
	(C) NÚMERO DE HEBRAS: simple	
	(D) CONFIGURACIÓN: Lineal	
65	(ii) TIPO DE MOLÉCULA: ADNc	
	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 78:	

CTGTCCGTTG GGTTTCCTTA CTCCT

25

(2) INFORMACIONES PARA LA SEC ID nº: 79:

5 (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 25 pares de bases

(B) TIPO: Nucleótido

(C) NÚMERO DE HEBRAS: simple

10 (D) CONFIGURACIÓN: Lineal

(ii) TIPO DE MOLÉCULA: ADNc

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 79:

15 AGGAGTAAGG AAACCCAACG GACAG

25

(2) INFORMACIONES PARA LA SEC ID nº: 80:

20 (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 25 pares de bases

(B) TIPO: Nucleótido

(C) NÚMERO DE HEBRAS: simple

25 (D) CONFIGURACIÓN: Lineal

(ii) TIPO DE MOLÉCULA: ADNc

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 80:

30 TGTATATAAT GGTCTGGCTA TTGGG

25

(2) INFORMACIONES PARA LA SEC ID nº: 81:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

35 (A) LONGITUD: 25 pares de bases

(B) TIPO: Nucleótido

(C) NÚMERO DE HEBRAS: simple

(D) CONFIGURACIÓN: Lineal

40 (ii) TIPO DE MOLÉCULA: ADNc

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 81:

AGGAGTAAGG AAACCCAACG GACAG

25

45 (2) INFORMACIONES PARA LA SEC ID nº: 82:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 26 pares de bases

50 (B) TIPO: Nucleótido

(C) NÚMERO DE HEBRAS: simple

(D) CONFIGURACIÓN: Lineal

(ii) TIPO DE MOLÉCULA: ADNc

55 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 82:

TTCGGCAGAA ACCTGTTATG CCAAGG

26

(2) INFORMACIONES PARA LA SEC ID nº: 83:

60 (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 22 pares de bases

(B) TIPO: Nucleótido

65 (C) NÚMERO DE HEBRAS: simple

(D) CONFIGURACIÓN: Lineal

	(ii) TIPO DE MOLÉCULA: ADNc	
	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 83:	
5	CTCGATTCT TGCTGGGCCT TA	22
	(2) INFORMACIONES PARA LA SEC ID nº: 84:	
10	(i) CARACTERÍSTICAS DE LA SECUENCIA:	
	(A) LONGITUD: 20 pares de bases	
	(B) TIPO: Nucleótido	
	(C) NÚMERO DE HEBRAS: simple	
	(D) CONFIGURACIÓN: Lineal	
15	(ii) TIPO DE MOLÉCULA: ADNc	
	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 84:	
	GTTGATTCCC TCCTCAAGCA	20
20	(2) INFORMACIONES PARA LA SEC ID nº: 85:	
	(i) CARACTERÍSTICAS DE LA SECUENCIA:	
25	(A) LONGITUD: 20 pares de bases	
	(B) TIPO: Nucleótido	
	(C) NÚMERO DE HEBRAS: simple	
	(D) CONFIGURACIÓN: Lineal	
30	(ii) TIPO DE MOLÉCULA: ADNc	
	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 85:	
	CTCTACCAAT CAGCATGTGG	20
35	(2) INFORMACIONES PARA LA SEC ID nº: 86:	
	(i) CARACTERÍSTICAS DE LA SECUENCIA:	
40	(A) LONGITUD: 19 pares de bases	
	(B) TIPO: Nucleótido	
	(C) NÚMERO DE HEBRAS: simple	
	(D) CONFIGURACIÓN: Lineal	
45	(ii) TIPO DE MOLÉCULA: ADNc	
	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 86:	
	TGTTCTCTT GGTCCTAT	19
	(2) INFORMACIONES PARA LA SEC ID nº: 87:	
50	(i) CARACTERÍSTICAS DE LA SECUENCIA:	
	(A) LONGITUD: 433 restos de aminoácidos	
	(B) TIPO: aminoácidos	
55	(ii) TIPO DE MOLÉCULA: péptido	
	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 87:	

```

Met Ala Thr Ala Thr Gly Thr Gly Ile Ala Gly Leu Ser Thr Ser Leu
1      5      10      15
Ser Tyr Tyr His Thr Leu Ser Lys Asn Phe Ser Asp Ser Leu Gln Glu
20      25      30
Ile Met Lys Ser Ile Leu Thr Leu Gln Ser Gln Leu Asp Ser Leu Ala
35      40      45
Ala Met Thr Leu Gln Asn Arg Arg Gly Pro His Leu Leu Thr Ala Glu
50      55      60
Lys Gly Gly Leu Cys Thr Phe Leu Gly Glu Glu Cys Cys Phe Tyr Thr
65      70      75      80
Asn Gln Ser Gly Ile Val Arg Asp Ala Thr Trp His Leu Gln Glu Arg
85      90      95
Ala Ser Asp Ile Arg Gln Cys Leu Ser Asn Ser Tyr Thr Asn Leu Trp
100     105     110
Ser Trp Ala Thr Trp Leu Leu Pro Phe Leu Gly Pro Met Ala Ala Ile
115     120     125
Leu Leu Leu Leu Thr Phe Gly Pro Cys Ile Phe Lys Leu Leu Val Lys
130     135     140
Phe Val Ser Ser Arg Ile Glu Ala Ile Lys Leu Gln Met Val Leu Gln
145     150     155     160
Met Glu Pro Gln Met Ser Ser Thr Asn Asn Phe Tyr Gln Gly Pro Leu
165     170     175
Glu Arg Ser Thr Gly Thr Ser Thr Ser Leu Glu Ile Pro Leu Trp Lys
180     185     190
Thr Leu Gln Leu Gln Gly Pro Phe Phe Ala Pro Ile Gln Gln Glu Val
195     200     205
Ala Arg Ala Val Ile Gly Gln Ile Pro Asn Ser Ser Trp Gly Val Leu
210     215     220
Phe Arg Gly Gly Ile Glu Glu Val Thr Ala Cys Trp Gln Pro His Ser
225     230     235     240
Pro Arg Trp Xaa Ser Val Pro Pro Gln Pro Trp Cys Pro Leu Trp Pro
245     250     255
Cys Leu Arg Ser Pro Ser Ala Cys His Cys Thr Val Gly Ala Ser Phe
260     265     270
Trp Ala Gly Gln Gly Arg Ser Gln Leu Pro Gln Leu Ala Gly Arg Tyr
275     280     285
Gly Gly Arg Asp Ala Gly Gly Asn Gln Gly Cys Ala Trp Arg Leu Arg
290     295     300
Ala Ser Met Ser Ser Arg Trp Ala Trp Ala Arg Ala Pro His Ser
305     310     315     320
Gly Ser Glu Gly Leu Ser Thr Trp Ala Arg Gln Met Leu Cys Ser Thr
325     330     335
Ser Ser Leu Gly Leu Ser Cys Leu Pro Arg Gly Ala Gly Leu Arg Glu
340     345     350
Xaa Ala Ala Cys Pro Cys Leu Ser Pro Pro Pro Arg Arg Gly Phe Leu
355     360     365
His Ser Pro Ser Phe Pro Asp Lys His His Pro Leu Ser Thr Val Pro
370     375     380
Ser Pro Ile Asn His Pro Arg Val Glu Glu Cys Gly His Thr Ala Arg
385     390     395     400

Asp Trp Gln Ala Val Pro Leu Ala Ala Leu Val Arg Asp Pro Leu Arg
405     410     415
Glu Ala Ser Trp Ala Pro Glu Ser Gly Gly Asp Leu Glu Asn Leu Tyr
420     425     430
Val
433

```

(2) INFORMACIONES PARA LA SEC ID nº: 88:

5

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 693 pares de bases
 (B) TIPO: Nucleótido
 (C) NÚMERO DE HEBRAS: simple
 (D) CONFIGURACIÓN: Lineal

10

(ii) TIPO DE MOLÉCULA: ADNc

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 88:

15

```

CTTCCCCAAC TAATAAGGAC CCCCCTTTCA ACCCAAACAG TCCAAAAGGA CATAGACAAA 60
GGAGTAAACA ATGAACCAAA GAGTGCCAAT ATTCCCTGGT TATGCACCCT CCAAGCGGTG 120
GGAGAAGAAT TCGGCCCAGC CAGAGTGCAT GTACCTTTTT CTCTCTCACA CTGAAGCAA 180
ATTAAAATAG ACNTAGGTNA ATTNTCAGAT AGCCCTGATG GYTATATTGA TGTTTTACAA 240
GGATTAGGAC AATCCTTTGA TCTGACATGG AGAGATATAA TATTACTGCT AAATCAGACG 300
CTAACCTCAA ATGAGAGAAG TGCTGCCATA ACTGGAGCCC GAGAGTTTGG CAATCTCTGG 360
TATCTCAGTC AGGTCAATGA TAGGATGACA ACGGAGGAAA GAGAACGATT CCCCACAGGG 420
CAGCAGGCAG TTCCAGTGT AGCTCCTCAT TGGGACACAG AATCAGAACA TGGAGATTGG 480
TGCCGCAGAC ATTTACTAAC TTGCGTGCTA GAAGGACTAA GGAAACTAG GAAGACTATG 540
AATTATTCAA TGATGTCCAC TATAACACAG GGGAAAGGAA GAAAATCCTA CTGCCTTTCT 600
GGAGAGACTA AGGGAGGCAT TGAGGAAGCA TACCAGGCAA GTGGACATTG GAGGCTCTGG 660
AAAAGGGAAG ACTTGGGCAA ATTGAATGCC TAA 693

```

(2) INFORMACIONES PARA LA SEC ID nº: 89:

5 (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 1577 pares de bases
 (B) TIPO: Nucleótido
 (C) NÚMERO DE HEBRAS: simple
 (D) CONFIGURACIÓN: Lineal

10

(ii) TIPO DE MOLÉCULA: ADNc

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 89:

```

AACTTGCGTG CTAGAAGGAC TAAGGAAAAC TAGGAAGACT ATGAATTATT CAATGATGTC 60
CACTATAACA CAGGGGAAAG GAAGAAAATC CTACTGCCTT TCTGGAGAGA CTAAGGGAGG 120
CATTGAGGAA GCATACCAGG CAAGTGGACA TTGGAGGCTC TGGAAAAGGG AAAAGTTGGG 180
CAAAATTGAAT GCCTAATAGG GCTTGCTTCC AGTGCAGTCT ACAAGGACGC TTTAGAAAAG 240
ATTGTCCAAG TAGAATAAAG CCGCCCCCTCG TCCATGCCCTT TTATGTCAAG GGAATCACTG 300
GAAGGCCTAC TGCCCCAGGG GACGAAGGTC CTCTGAGTCA GAAGCCACTA ACCTGATGAT 360
CCAGCAGCAG GACTGAGGGT GCCCGGGGCA AGTGCCAGCC CATGCCATCA CCCTCAGAGC 420
CCCGGGTATG TTTGACCATT GAGAGCCAGG AAGTTAACTG TCTCCTGGAC ACTGGCGCAG 480
CCTTCTCAGT CTTACTTTCC TGTCCCAGAC AATTGTCTCT CAGATCTGTC ACTATCCGAG 540
GGGTCTTAAG ACAGCCAGTC ACTACATACT TCTCTCAGCC ACTAAGTTGT GACTGGGGAA 600
TTTACTCTT TTACATGCT TTTCTAATTA TGCTGAAAG CCCCACTCCC TTGTTAGGGA 660
GAGACATTTT AGCAAAAGCA GGGGCCATTA TACACCTGAA CATAGGAAAA GGAATACCCA 720
TTTGCTGTCC CCTGCTTGAG GAAGGAATTA ATCCTGAAGT CTGGGCAATA GAAGGACAAT 780
ATGGACAAGC AAGCAATGCC CGTCCTGTTC AAGTTAACT AAAGGATTCT GCCTCCTTTC 840
CCTACCAAGG GAAGTACCTT CTTAGACCCG AGGCCCTACA AGGACTCAA AGATTGTTAA 900
GGACCTAAAG GCCCAAGGCC TAGTAAACC ATGCAGTAGC CCCTGCAATA CTCCAATTTT 960
AGGAGTAAGG AAACCCAACG GACAGTGGAG GTTAGTGCAA GATCTCAGGA TTATTAATGA 1020
GGCTGTTTTT CCTCTATACC CAGCTGTATC TAGCCCTTAT ACTCTGCTTT CCCTAATACC 1080

```

```

AGAGGAAGCA GAGTAGTTTA CAGTCCTGGA CCTTAAGGAT GCCTCTTTCT GCATCCCTGT 1140
ACATCCTGAT TCTCAATTCT TGTTGTCTT TGAAGATCCT TTGAACCCAA TGTCTCAATT 1200
CACCTGGACT GTTTTACCCC AGGGGTTCCG GGATAGCCCC CATCTATTTG GCCAGGCATT 1260
AGCCCCAAGC TTGAGCCAAT TCTCATACCT GGACATCTTG TCCTTCGGTA TGGGATGATT 1320
TAATTTTATG CAACCGTTCA GAAACCTTGT GCCATCAAGC CACCCAAGCG TTCTTAAATT 1380
TCCTCACTCC GTGTGGCTAC AAGGTTTCCA AACCAGAGGC TCAGCTCTGC TCACAGCAGG 1440
TTAAATACTT AGGGTTAAAA TTATCCAAAG GCACCAAGGC CCTCTGTGAG GAATGTATCC 1500
AACCTGTACT GGCTTATCTT CATCCCAAAA CCCTAAAGCA ACTAAGAAGG TCCTTGGCAT 1560
AACAGGTTTC TGCCGAA 1577

```

15

(2) INFORMACIONES PARA LA SEC ID nº: 90:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 182 restos de aminoácidos
 (B) TIPO: aminoácidos

20

(ii) TIPO DE MOLÉCULA: péptido

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 90:

25

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

(2) INFORMACIONES PARA LA SEC ID nº: 91:

5 (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 36 pares de bases
- (B) TIPO: Nucleótido
- (C) NÚMERO DE HEBRAS: simple
- (D) CONFIGURACIÓN: Lineal

(ii) TIPO DE MOLÉCULA: ADNc

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 91:

15 AGATCTGCAG AATTCGATAT CACCCCCCCC CCCCCC

36

(2) INFORMACIONES PARA LA SEC ID nº: 92:

20 (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 22 pares de bases
- (B) TIPO: Nucleótido
- (C) NÚMERO DE HEBRAS: simple
- (D) CONFIGURACIÓN: Lineal

(ii) TIPO DE MOLÉCULA: ADNc

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID nº: 92:

30 AGATCTGCAG AATTCGATAT CA

22

REIVINDICACIONES

1. Polipéptido antigénico reconocido por sueros de pacientes infectados por un retrovirus cuyo genoma comprende una secuencia nucleotídica seleccionada de entre SEC ID nº 1 y SEC ID nº 89, y/o en los cuales se ha reactivado dicho retrovirus, cuya secuencia peptídica consiste en una secuencia seleccionada de entre SEC ID nº 39, SEC ID nº 41 a SEC ID nº 44 y SEC ID nº 63.
2. Polipéptido según la reivindicación 1, caracterizado porque posee una secuencia peptídica que consiste en una secuencia seleccionada de entre SEC ID nº 39 y SEC ID nº 63.
3. Secuencia nucleica que codifica para un polipéptido según la reivindicación 1 ó 2, seleccionada de entre SEC ID nº 40, SEC ID nº 62, sus fragmentos y las secuencias complementarias de éstos.
4. Anticuerpo mono- o policlonal dirigido contra un agente patológico y/o infeccioso asociado a la esclerosis en placas, caracterizado porque se obtiene mediante reacción inmunológica de un organismo humano o animal, con un agente inmunógeno constituido por un polipéptido según la reivindicación 1 ó 2.
5. Agente reactivo de detección de un agente patológico y/o infeccioso asociado a la esclerosis en placas, o de una exposición a dicho agente, caracterizado porque comprende, a título de sustancia reactiva, un polipéptido según la reivindicación 1 ó 2, o un anticuerpo según la reivindicación 4.
6. Composición diagnóstica, profiláctica o terapéutica, que comprende un polipéptido según la reivindicación 1 ó 2, o un anticuerpo según la reivindicación 4.
7. Composición inmunoterapéutica activa, en particular composición vacunal, según la reivindicación 6.
8. Procedimiento para detectar la presencia o la exposición a un agente patológico y/o infeccioso asociado a la esclerosis en placas, en una muestra biológica, caracterizado porque se pone en contacto dicha muestra con un polipéptido según la reivindicación 1 ó 2, o un anticuerpo según la reivindicación 4.
9. Dispositivo de detección de la presencia o de la exposición a un agente patológico y/o infeccioso asociado a la esclerosis en placas, en una muestra biológica, comprendiendo dicho dispositivo un agente reactivo según la reivindicación 5, que consiste en un polipéptido, soportado por un soporte sólido, inmunológicamente compatible con dicho agente reactivo, caracterizado porque comprende un medio de puesta en contacto de dicha muestra biológica con dicho agente reactivo, en condiciones que permiten una eventual reacción inmunológica, y unos medios de detección del complejo inmune formado con dicho agente reactivo.
10. Procedimiento de detección de anticuerpo dirigido contra un agente patológico y/o infeccioso asociado a la esclerosis en placas, en una muestra biológica, caracterizado porque se pone en contacto dicha muestra y un agente reactivo según la reivindicación 5, que consiste en un anticuerpo, en condiciones que permiten una eventual reacción inmunológica, y se detecta a continuación la presencia de un complejo inmune formado con dicho agente reactivo.
11. Procedimiento según la reivindicación 10, caracterizado porque la muestra es una muestra de sangre o de líquido cefalorraquídeo.
12. Utilización de una región inmunodominante seleccionada de entre SEC ID nº 41 a SEC ID nº 44, para preparar unos péptidos antigénicos reconocidos por sueros de pacientes infectados por un retrovirus cuyo genoma comprende una secuencia nucleotídica seleccionada de entre SEC ID nº 1 y SEC ID nº 89, y/o en los cuales se ha reactivado dicho retrovirus.

FIG. 1

Consenso GTTCAGGGAT AGCCCTCATC TCTTTGGTCA GGTACTGGOC CAAGATCTAG 50

Consenso GCCACTTCTC AGGTCCAGSN ACTCTGTGTC TTGAG 85

SEC ID n° 3 (POL MSRV-1B)

Consenso GTTCAGGGAT AGCCCCATC TATTGGGCA GGTACTAGCT CAACTACTGA 50

Consenso GGCAGTCTC ATACTGGAC ATCTGTGTC TTGGT 85

SEC ID n° 4 (POL MSRV-1B)

Consenso GTTCAGGAT AGCCCCATC TATTGGGCA GGTACTAGCT CAAGACTGA 50

Consenso GYCAATCTC ATACTGGAC ACTCTGTGTC TTYRG 85

SEC ID n° 5 (POL MSRV-1B)

Consenso GTTCAGGGAT AGCTCCATC TATTGGGCT GGCATTAOC CGAGACTTA 50

Consenso GGCAGTCTY ATAAGTGGC ACTCTGTGTC TTGG 85

SEC ID n° 6 (POL MSRV-1B)

Consenso GTGTGGCAC AGGGGTTAR RGATANCY CATCTMTTG GYWRGYXT

Consenso RRCYCRKAY YIRRGYAVT TCTYAKRYSY RGSNAYTCTB KYOCTTYRGT

Consenso ACATGGATGA C

SEC ID n° 7 (POL MSRV-1B)

FIG.2

CONSENSO A**SEC ID nº 3**

GTTTAGGGATAGCCC TCATCTCTTGGTCA GGTACTGGCCCAAGA TCTAGGCCACTTCTC 60
 V G P S S L W S G T G P R S R P L L
 F R D S P H L F G Q V L A Q D L G H F S
 L G I A L I S L V R Y W P K I A T S Q

AGGTCCAGGCACTCT GTTCCTTCAG 85
 R S R H S V P S
 G P G T L F L Q
 V Q A L C S F

CONSENSO B**SEC ID nº 4**

GTTCAGGGATAGCCC CCATCTATTTGGCCA GGCCTAGCTCAATA CTTGAGCCAGTTCTC 60
 V Q G . P P S I W P G T S S I L E P V L
 F R D S P H L F G Q A L A Q Y L S Q F S
 S G I A P I Y L A R H . L N T . A S S H

ATACCTGGACACTCT TGTCTTCGGT 86
 I P G H S C P S
 Y L D T L V L R
 T W T L L S F G

CONSENSO C**SEC ID nº 5**

GTTCAGGGATAGCCC CCATCTATTTGGCCA GGCATTAGCCCAAGA CTTGAGTCAATTCTC 60
 V Q G . P P S I W P G I S P R L E S I L
 F R D S P H L F G Q A L A Q D L S Q F S
 S G I A P I Y L A R H . P K T . V N S H

ATACCTGGACACTCT TGTCTTCAG 85
 I P G H S C P S
 Y L D T L V L Q
 T W T L L S F

CONSENSO D**SEC ID nº 6**

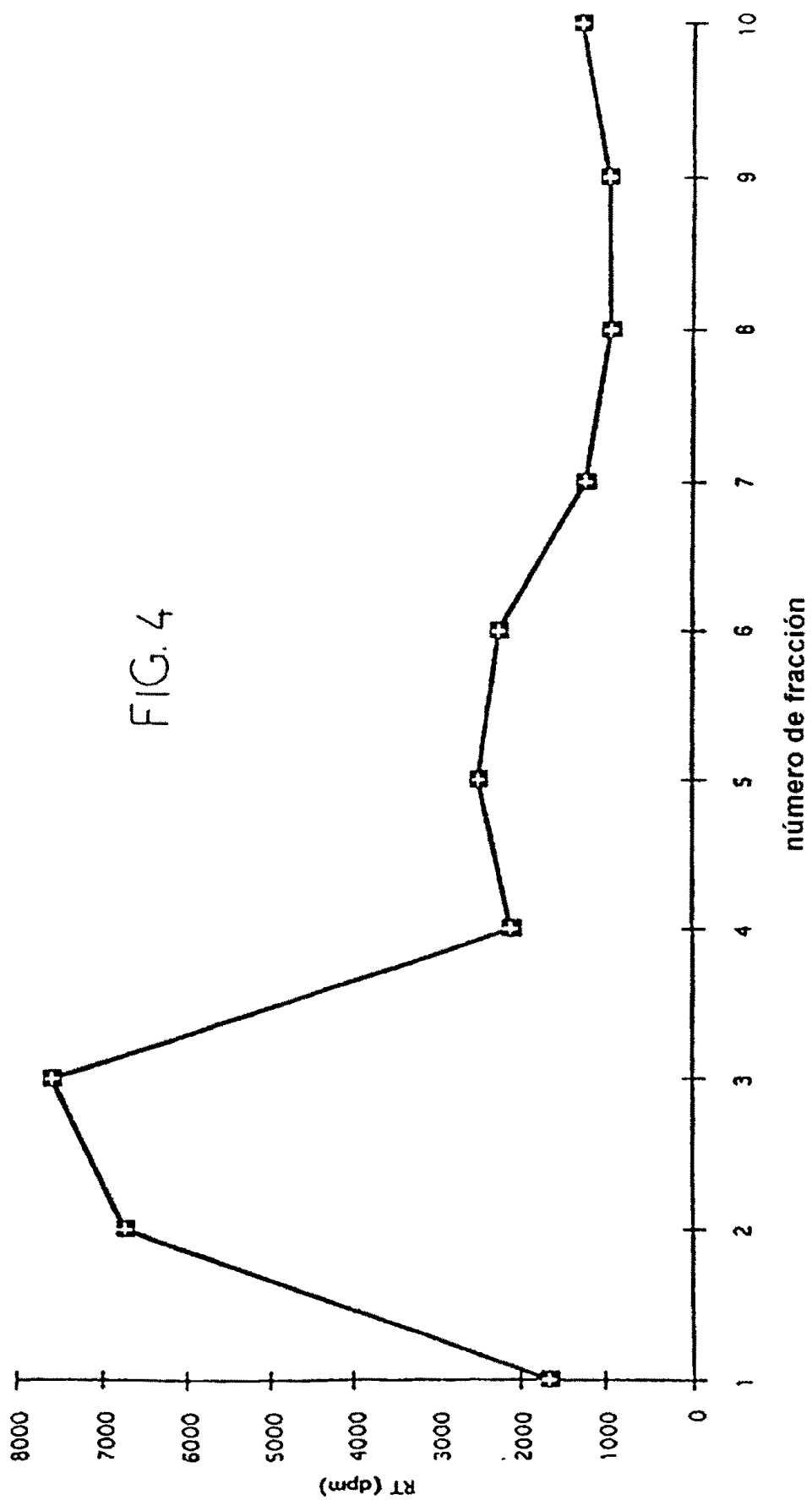
GTTCAGGGATAGCTC CCATCTATTTGGCCT GGCATTAACCCGAGA CTTAAGCCAGTTCTC 60
 V Q G . L P S I W P G I N P R L K P V L
 F R D S S H L F G L A L T R D L S Q F S
 S G I A P I Y L A W H . P E T . A S S H

ATACGTGGACACTCT TGTCTTTGG 85
 I R G H S C P L
 Y V D T L V L W
 T L L S F

FIG. 3

Consenso	TTGGATCCAG TGYTGCCACA GGGCGCTGAA GGCATCCCG TCCAGTTGCC	53
Consenso	GGATGCGGOC TATAGCCTCT ACGTGCATGA CCCTCTGAAG CTTCAG	96

SEC ID n° 11



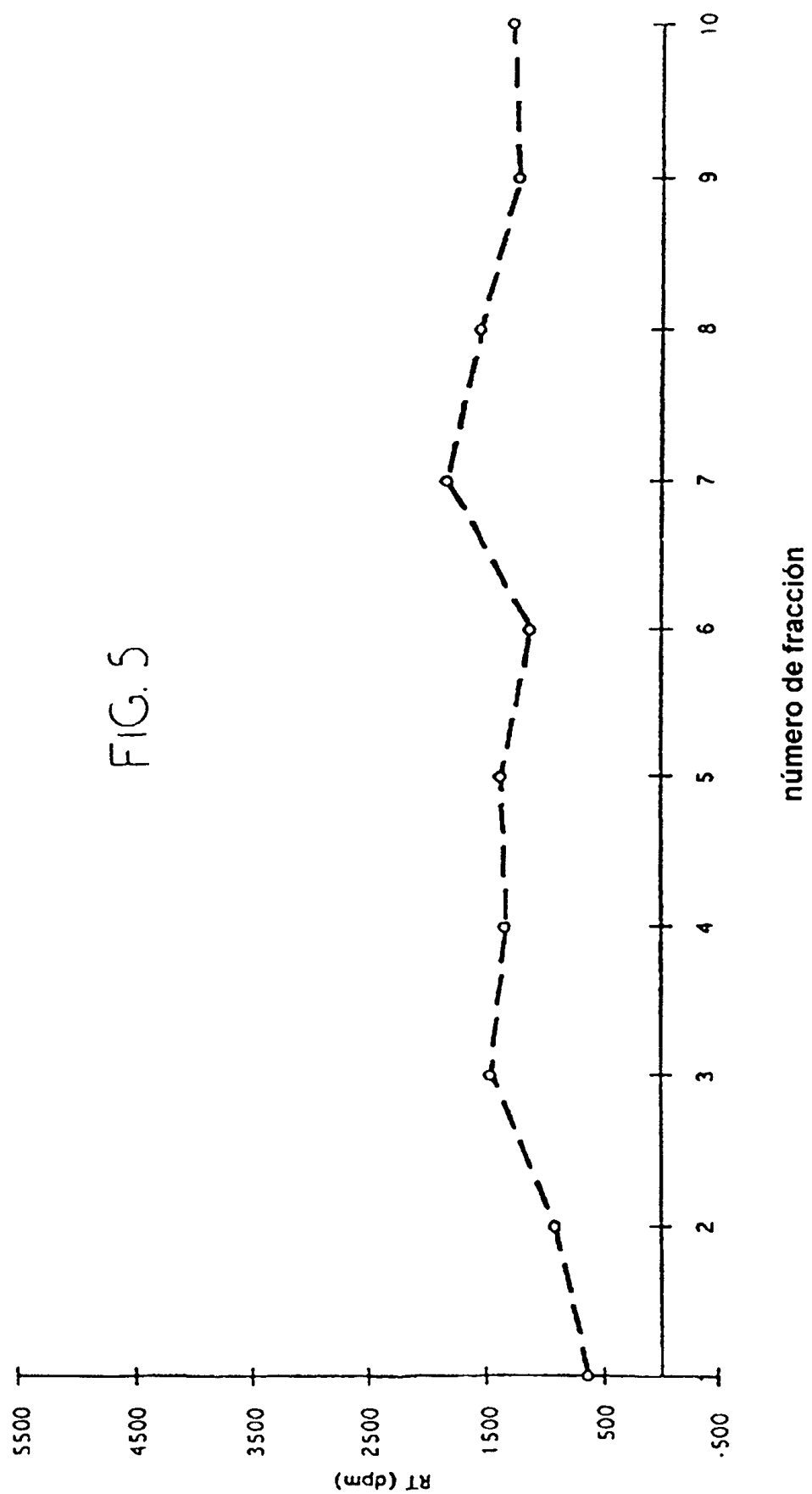


FIG. 6

CAAGCCACCC AAGAACTCTT AAATTTCTCT ACTACCTGTG GCTACAAGGT	50
TTCCAAACCA AAGGCTCAGC TCTGCTCACA GGAGATTAGA TACTTAGGGT	100
TAAAATTATC CAAAGGCACC AGGGGCTCA GTGAGGAACG TATCCAGCCT	150
ATACTGGGTT ATCTTCATCC CAAAACCTA AAGCAACTAA GAGGGTTCTT	200
TAGCATGATC AGGTTTCTGC CGAAAACAAG ATTCCCAGGT ACAACCAAAA	250
TAGCCAGACC ATTATATACA CTAATTAAGG AAACTCAGAA AGCCAATACC	300
TATTTAGTAA GATGGACACC TAAACAGAAG GCTTTCCAGG CCTTAAAGAA	350
GGCCCTAACC CAAGCCCCAG TGTTTCAGCTT GCCAACAGGG CAAGATTTTT	400
CTTTATATGG CACAGAAAAA ACAGGAATCG CTCTAGGAGT CCTTACACAG	450
GTCGAGGGA TGAGCTTGCA ACCCGTGGCA TACCTGAATA AGGAAATTGA	500
TGTAGTGGCA AAGGGTTGGC CTCATNGTTT ATGGGTAAATG GNGGCAGTAG	550
CAGTCTNAGT ATCTGAAGCA GTTAAAATAA TACAGGGAAG AGATCTINCT	600
GIGTGGACAT CTCATGATGT GAACGGCATA CTCACTGCTA AAGGAGACTT	650
GIGGTTGTCA GACAACCATT TACTTAANTA TCAGGCTCTA TTACTIONTGAAG	700
AGCCAGTGCT GNGACTGGC ACTTGTGCAA CTCTTAAACC C	741

SEC ID n° 9 (PSJ 17)

TCAGGGATAGCCCCCATCTATTTGGGCAGGCATTAGCCCAAGACTTGAGTC
AATTCTCATACCTGGACACTCTTGTCCTTCAGTACATGGATGATTTACTTT
TAGTCGCCCCGTT CAGAAACCTTGTGCCATCAAGCCACCCAAGA ACTCTTAA
CTTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCGGCTCT
GCTCACAGGAGATTAGATACTNAGGGCTAAAATTATCCAAAGGCACCAGG
GCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATCCCAAA
ACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAAA
ACAGATTCCCAGGTACASCCCAATAGCCAGACCATTATATACACTAATTA
NGGAAACTCAGAAAGCCAATACCTATTTAGTAAGATGGACACCTACAGAA
GTGGCTTTCAGGOCCTAAAGAAGGCCCTAACCCAAGCCCCAGTGTT CAGC
TTGCCAACAGGGCAAGATTTTTCTTTATATGCCACAGAAAAAACAGGAAT
AGCTCTAGGAGTCCTTACGCAGGTCTCAGGGATGAGCTTGCAACCCGTGGT
ATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTT

SEC ID n° 8 (MOO3-POO4)

FIG. 7

FIG. 9a

SEC ID n° 1 (MSRV-1 pol)

```

      10      20      30      40      50      60      70
CCC TTT GGC ACT ACA TCA ATT TTA GGA GTA AGG AAA CCC AAC GGA CAG TGG AGG TTA GTG CAA GAA CTC AGC
P F A T T S I L G V R K P M G Q W R L V Q E L R>
      80      90      100      110      120      130      140
ATT ATC AAT GAG GCT GTT GTT CCT CTA TAC CCA GCT GTA OCT AAC OCT TAT ACA GTG CTT TCC CAA ATA CCA
I I N E A V V P L Y P A V P H P Y T V L S Q I P>
      150      160      170      180      190      200      210
GAG GAA GCA GAG TGG TTT ACA GTC CTG GAC CTT AAG GAT GGC TTT TTC TGC ATC OCT GTA OCT CTT GAC TCT
E E A E W F T V L D L K D A F F C I P V R P D S>
      220      230      240      250      260      270      280
CAA TTC TTG TTT GGC TTT GAA GAT CCT TTG AAC CCA AGC TCT CAA CTC ACC TGG ACT GTT TTA CCC CAA GGG
Q F L F A F E D P L N P T S Q L T M T V L P Q G>
      290      300      310      320      330      340      350      360
TTC AGG GAT AGC CCC CAT CTA TTT GGC CAG GCA TTA GGC CAA GAC TTG AGT CAA TTC TCA TAC CTG GAC ACT
F R D S P H L F G O A L A O D L S Q F S Y L D T>
      370      380      390      400      410      420      430
CTT GTC CTT CAG TAC ATG GAT GAT TTA CTT TTA GTC GGC GGT TCA GAA ACC TTG TGC CAT CAA GGC ACC CAA
L V L O Y M D D L L V A R S E T L C H Q A T Q>
      440      450      460      470      480      490      500
GAA CTC TTA ACT TTC CTC ACT ACC TGT GGC TAC AAG GTT TCC AAA CCA AAG GCT CCG CTC TGC TCA CNG GAG
E L L T F L T T C G Y K V S K P K A R L C S Q G E>
      510      520      530      540      550      560      570
ATT AGA TAC TAA GGG CTA AAA TTA TCC AAA GGC ACC AGC GGC CTC AGT GAG GAA OCT ATC CAG OCT ATA CTG
I R Y X G L K L S K G T R A L S E E R I C Q P T L>
      580      590      600      610      620      630      640
GCT TAT OCT CAT CCC AAA ACC CTA AAG CAA CTA ACA AAG TTC CTT GGC ATA ACA GGT TTC TCC CGA AAA CAG
A Y P H P K T L K U L R G F L L G I T G F C R K Q>
      650      660      670      680      690      700      710      720
ATT CCC AGG TAC ASC CCA ATA GGC AGA CCA TTA TAT ACA CTA ATT AAG GAA ACT CAG AAA GGC AAT ACC TAT
I P R Y X P I A K P L V T L I X E T Q K A H T Y>
      730      740      750      760      770      780      790
TTA GTA AGA TGG ACA CCT ACA GAA GTG GCT TTC CAG GGC CTA AAG AAG GGC CTA ACC CAA GGC CCA GTG TTC
L V R W T P T E V A F Q A L K K A L T Q A P V F>
      800      810      820      830      840      850      860
AGC TTG CCA ACA GGG CAA GAT TTT TCT TTA TAT CCC ACA CAA AAA ACA CCA ATA OCT CTA GGA GTC CTT ACG
S L P T G Q D F S L Y A T E N T G I A L C V L T>
      870      880      890      900      910      920      930
CAG GTC TCA GGG ATG AGC TTG CAA CCC GTG GTA TAC CTG AGT AAG GAA ATT CAT CTA GTC CCA AAG GGT TGG
Q V J G H S L O P V V V V I S K F I D V V A K G W>
      940      950      960      970      980      990      1000
CCT CAT NGT TTA TGG CTA ATG CAG GCA GTA GGA GTC TAA CTA TCT CAA GCA GTT AAA ATA ATA CNG GGA AGA
P H X I W V M X A V A V X V Y Z E A V K I I U G H>
      1010      1020      1030      1040      1050      1060      1070      1080
GAT CTT NCT GTC TGG ACA TCT CAT CAT GTC AAC GGC ATA CTC ACT OCT AAA GGA GAC TTG TGG TTG TCA GAC
D L X V W T J H U V N C I L T A K G D L M L S D>
      1090      1100      1110      1120      1130      1140      1150
AAC CAT TTA CTT AAT TAT CAG GCT CTA TTA CTT GAA GAG CCA GTG CTG NAA CTG CCG ACT TGT CCA ACT CTT
N H L A L E E D V L X L R T C A T L>
      AAA CCC
      X P>

```

FIG. 9b

AAC CAT TTA CTT AAN TAT CAG GCT CTA TTA CTT GAA GAG CCA GTG CTG NGA CTG CGC ACT TGT GCA ACT CTT
N H L L X Y Q A L L L E E P V L X L R T C A A T L
_a_a_a_a_a_a_ TRANSLATION OF MSRV-1 POL * (A) _a_a_a_a_a_a_a_a_>
AAA CCC
K P>
a>

SEC ID NO 1 (MSRV-1 pol.)

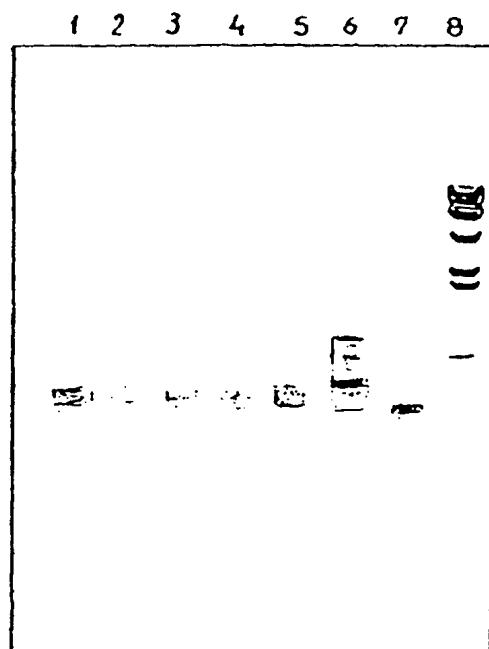


FIG. 10

FIG. 11

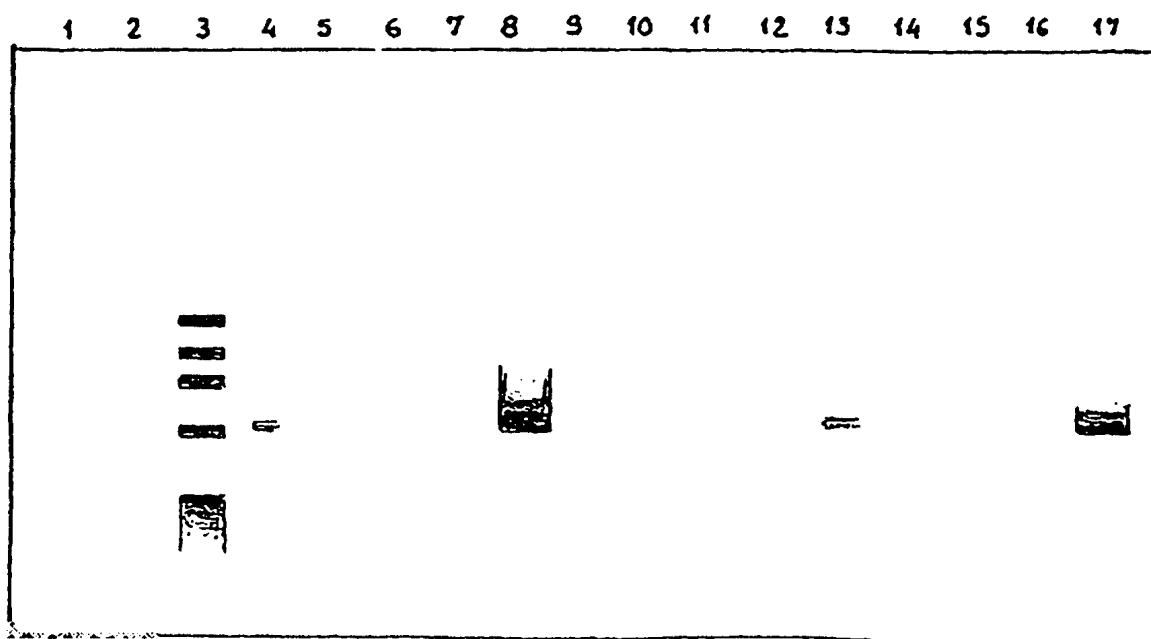


FIG. 12

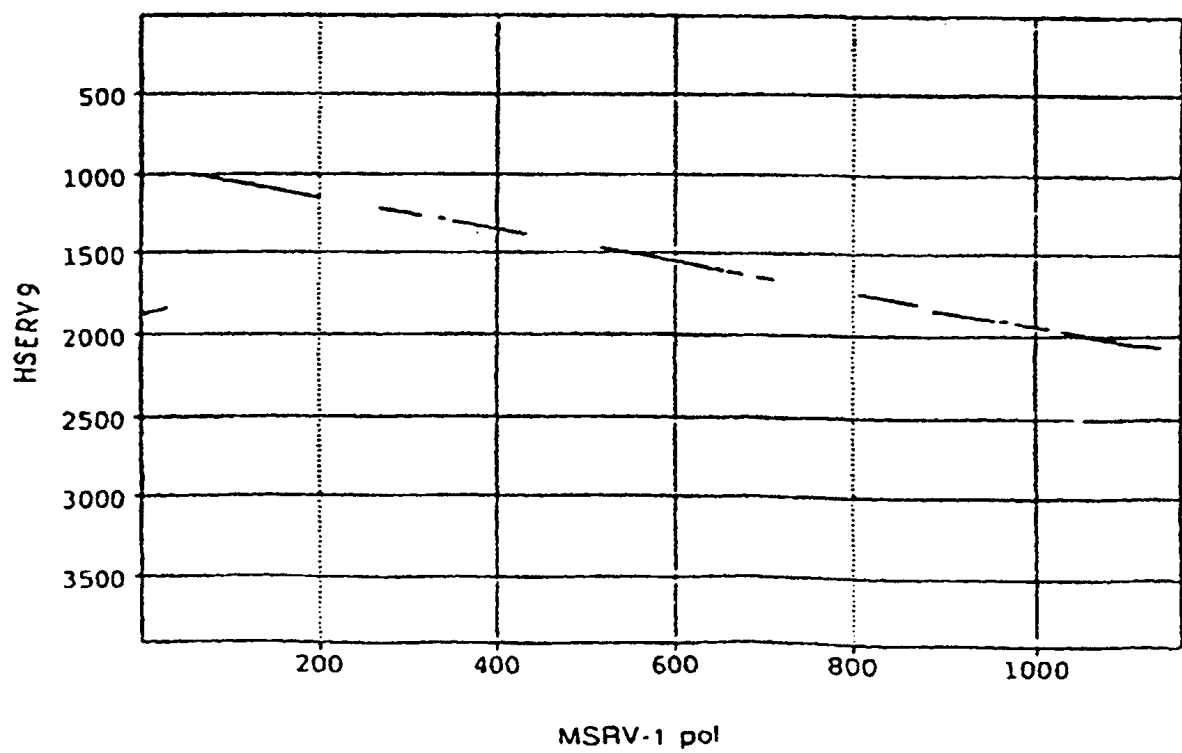


FIG. 13

SEC ID n° 46 (FBd3)

GTGCTGATTGGTGTATTTACAATCCTTTATCTAATCCGAAATGCCCATGTTG
CAATATGGAAAGAAAGGGAGTTCCTAACCTCTGGGGGAACCCCCATTAAA
TACCACAAGTAAATCATGGAGTTATTGCACACAGTGCAAAAAGTCAAGGA
GGTGGAAGTCTTACACTGCCAAAGCCATCAGAAAAGGGAAGAGGGGGAGAA
GAGCAGCATAAGTGGCTACAGAGGCAAGGAAAGACTAGCAGAAAGGAAA
GAGAGAAAGAGACAGAAAGTCAGAGAGAGAGAGAGGAAGAGACAGAGCA
CAAAGAGGGAGTCAGAGAGAGAGAGAGACAGAGAGTCAGAGAGAAGGAA
AGAGAGAGAGGAAGAGACAAAGAATGAATCAAACAGAGAGACAGAAAGT
CAGAGAGAGAGAGAGAGAGGAAGAGACAGAGAAAAAGAGGGGAGTCAGAA
AAAGAGAGACCAAAGAAGAAGTCCAAAGAGAAAGAAAGAGAGATGGAAG
TAGTAAAGGAAAAACAGTGTACCCTATTCTTTAAAAGCCGGGGTAAATTT
AAAACCTATAATTGATAACTGAAGGTCTTCTCTGTAACCCTGTAACACTCC
AATACCACCTTGTTGTCAAGTGTAAACAAGGGCGTAGCCCCAAAAGCACTG
AGGCCACTAACAAACCCATAGCCTTCCTATCAAAATTCCTTAACCCAGCAGG
TTTCCTAACAGGGGATCTAAATCTTAATTAATTACCATACAATGGTCCAAC
CAGACTTAGGAGGAATTCCCTTCAGGACGGGAAGATAGATGCTTCCTCCCA
GGCGATTAAAGGGAGAAAGACACAATGGGTATTTCAGTAAGTGCCAAGGGGA
ACACTTGTAGAAGCAAAGTTAGGAAAATTGCCAAATAATTGGTTTGCTCAA
GAGTTGTTTGCACTCAGCCAAACCTTGAAGTACTTGCAGAATCAGAAAGGA
GCCATCTATACCAATTCTAAGTTAATATGGACTGAAGGAGGTTTTATTAAT
ACCAAAGAGAAATTAATAATCCCAAACCTTATAAGGTTTTCAACCAAAGTAA
AGTTTGCTAAAAGTTAACAGCGTAACATGTATTATCCTACTACCACACACT
CTCAAAGGATTTCTCAGACAGTTTGCAAGAAATAATGATATCTATCCTTAC
TCTACAATCCCAAATAGACTCTTTGGCAGCAGTGACTCTCCAAAACCGTCA
AGGCCTAGACCTCCTCACTGCTGAGAAAGGAGGACTCTGCACCTTCTTAAG
GGAAGAGTGTTGTCTTTACACTAACCAGTCAGGGATAGTATGAGATGCTGC
CCGGCATTACAGAAAAAGGCTTCTGAAATCAGACAACGCCTTTCAAATTC
CTATACCAACCTCTGGAGTTGGGCAACATGGTTTCTTCCCTTTCTATGTCCC
ATGGCTGCCATCTTGCTATTACTCGCCTTTGGGCCCTGTATTTTAAACCTCC
TTGTCAAATTTGTTTCTTCTAGGATCGAGGCCATCAAGCTACAGATGGTCTT
ACAAATGGAACCCCAAATGAGCTCAACTATCAACTTCTACTGAGGACCCCT
AGACCAACCCCTGGCCCTTCACTGGCCTAAAGAGTTCCCGTCTGGAGGA
CACTACCACTGCAGGGCCCCATCTTTGCCCTATCCAGAAGGAAGTAGCTA
GAGCAGTCATTGCCCAATTCCCAAGAGCAGCTGGGGTGTCCCGTTTAGAGT
GGGGATTGAGAGGTGAAGCCAGCTGGACTTCTGGGTGCGGGTGGGGACTTG
GAGAACTTTTGTGTCTAGCTAAAGGATTGTAAATGCAACAATCAGTGCTCT
GTGTCTAGCTAAAGGATTGTAAATACACCAATCAGCAC

FIG. 14

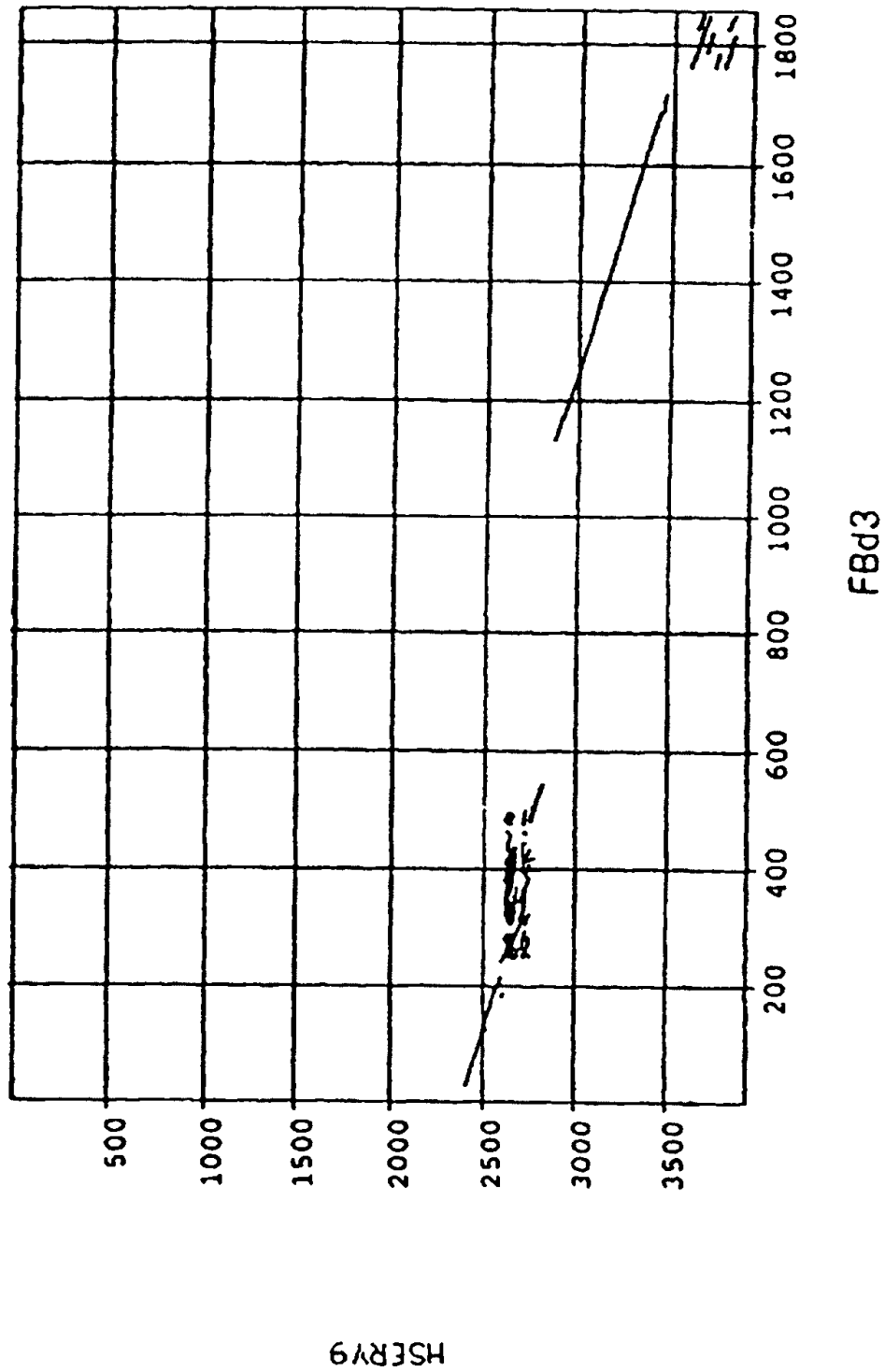


FIG. 15

SEC ID n° 51 (t pol)

GGCTGCTAAAGGAGACTTGTGGTTGTCAGACAATCGCCTACTTAGGTACCA
GGCCTTATTACTTGAGGGACTGGTGCTTCAGATGCGCACTTGTGCAGCTCT
TAACCCAAACTTATGCTGCCCAGAAGGATCTTTTAGAGGTCCCCTTAGCCA
ACCCTGACCTCAACCTATATATATACTGATGGAAGTTCGTTTGTAGAAAAG
GGATTACAAAGGGNAGGATATNCCATAGGTTAGTGATAAAGCAGTACTTG
AAAGTAAGCCTCTTCCCCCAGGGACCAGCGCCCCCGTTAGCAGAACTAGT
GGCACTGACCCCGAGCCTTAGAACTTGGAAAGGGAGGAGGATAAATGTGT
ATACAGATAGCAAGTATGCTTATCTAATCCGAAATGCCCATGTTG

SEC ID n° 52 (JLBc1)

TCAGGGATAGCCCCCATCTATTTGGTCAGGCACCTGGCCCAAGATCTAGGGA
 CATGCCACTTTTAAAGAGCCATTTCTCAAGTCCAGGTACTCTGGTCCITCGGT
 ATGTGGATGATTTACTTTTGGCTACCAGTTCAGTAGCCTCATGCCAGCAGG
 CTACTCTAGATCTCTTGAACCTTTCTAGCTAATCAAGGGTACAAGGCATCTA
 GGTTGAAGGCCAGCTTTGCCTACAGCAGGTCAAATATCTAGGCCTAATCT
 TAGCCAGAGGGGACCAGGGCACTCAGCAAGGAACAAATACAGCCTATACTG
 GCTTATCCTCACCCCTAAGACATTAACACAGTTGCGGGGGTTTCTTGGGAATC
 ACTGGCTTTTTGGTGACTATGGATTCCCAGATACAGCAAGATTGGCAGGCC
 CCTCTATACTGTAATCAAGGAGACTCACGAGGGCAAGTACTCATCTAGTAG
 AATGGGAACCTAGGGACAGAAACAGCCTTCAAAACCTTAAAGCAGGGCCCTA
 GTACAATCTCCAGCTTTAAGCCTTCCCACAGGACAAAACCTTCTCTTTATAC
 ATCACAGAGAGGGCAGAGATAGCTCTTGGTGTCTTATTACAGACTCATGGG
 ACTACCCCAACAACCAGTGGCACACCTAAGTAAGGAAATTGATGTAGTAGC
 AAAAGGCTGGCCTCACTGTTTATGGGTAGCTGTGGTGGTGGCTGTCTTAGT
 GTCAGAAGCTATCAAAATAATACAAGGAAAGGATCTCACTGTCTGGACTA
 CTCATGATGTAATGGCATACTAGGTGCCAAAAGAAGTTTATGGGTATCAGA
 CAACCACCTGCTTAGATACCAGGGACTACTCCTGGAGGATTGGGCTTCAAG
 TGGCTTTTTTGTGGCCTCAACCCTGCCACTTTTCTCCAGAGGATGGAGAG
 CCGCTTGAGCATGCTTGCCAACAGGTTGTAGGCCAGAATTATTCCACCCGA
 GATGATCTCTTAGAGTACCCTTAGCTAATCCTGACCTTAACCTATATACCA
 ATGGAAGTTCATTTGTGGAAAACGGGATATGAAGGGCAGGTTATGTCTATAG
 TTAGTGATGTAATCATACTTGCAAGTAAGCCTCTTACCCCAAGGGGCCAGCA
 CTCAGTTAGCAGAAGTAGTCACACTTACCTTAACCTTAGAACTGGGAAAGG
 GAAAAAGAATAAATATGTATACAGATAGTAAGTATGCTTATCTAATCCTAC
 ATGCCCATGCTGCAATATGGAAGGAAAGGGAGTTCCTAACCCTTGGGGGA
 ACCCCCATTAATAACCACAAGGYAAATCATGGAGTTATTGCACGCAGTGC
 AAAAACTCAAGGAGGTGGCAGTCTTACACTGCCGAAGCYATCAAAAAGGG
 GAAGGAGAGGGGAGAACAGCAGCATAAGTGGTTGGCAGAGGCAGTGAAA
 GACCAGCAGAGAGAAGGAGAGAGACAACGTCAACGACAGAAGGAAAGAA
 GAGGAGGAGACAGAGAGGAAGAGACAGAGAGACAGTTAGTCCAAGAGAG
 AGACAGAGAGAGGAAGAGACAGACAGAAAAGTCCAAGAGAGAAGGAAAGA
 GAGGAAGAGACCAAGGAGTCCNAGAGAGAGAAAAGAGATAGAAGTAGTAA
 AGAAAAAACATTGTACCCTATTCTTTAAAAGCCGGGGTATATTAAAAACC
 TATAATTGATAATTGAGTTCTTGACCCCTCCTCCAGGGGATYGCTGGGAGG
 AAACCCTCAACCGATATGTGAAAATTGTGGGTGCTCCCTATGTCTCAATTA
 CCAGCCAATACCCCTTGTTTTTAGTGTGAACGAGGGTGTAGAGCGCAGAC
 AGGGAGACCTCTGACAATCCATACCCTTCTATCCAAAATCCTTAACCCAG
 CAGGTTTTCTAAAAGGGGATCTAAATCTTAATTAATTACCATACAAAGGTC
 AAACCAGATCTAGGAGGAACTTCTTCAGGACAGGATGATAGATGGTTCCT
 CCCAGGCGATTAAAGAAAATAAAAAGACACATGGGCAGCCAGTAAGTGAT
 AAGGGAACACTAGTAGAAGCAGTTAGGAGAAGTTGCCTAATAATTGGTCT
 ACTCCAAATGTGTGAGTTGTTCCGCACTCAGCCCAAATCTTAAAGTACTTAC
 AGAATTAGGGAGGAGCCATTTACACCAATTCTAAGTTAATATGGACTGGAT
 GAGGTTTTATTAATAGCGAAGGAGAATTAAATCCTAAACTNACAAGGTTTT
 CAACTAAAGTAAATTTTACTAAAAGCTAACAGTGTAACATGCATTATCCTA
 CTACAACACACTCTCANAGGATTCTCAGACAGTTTACAAGAAATAACAA
 AATCTATCTGGTAAGGATAGTAACTACAATCCCAAATACATTCTTTGGCAG
 CAGTGACTCTC

FIG. 16

SEC ID n° 53 (JLBc2)

TCAGGGATAGCCCCC.TCTATTTGATCAGGCACTAGCCCAAGATCTAGGCC
 ACTTCTGAAGTCCAGGCAATTCTAGTCCTTCAGTATGTGGATGATTTACTTTT
 GGCTACCAAGTTTGGAAGCCTCATGCCAGCAGGCTACTTGAGATCTCTTGAA
 CTTTCTAGCTAATCAAGGGTGTATGGCATCTAAATTGAAAGTCCAGCTCTG
 CCTACAACAAGTCAAATATCTAGGCCTAATCTTAGATAGAAGAACCAGGG
 CCCTCAGCAAGGAATGAATAAAGCCTATGCTGGCTTATCGGCACCCTAAGA
 CATTAACAATTTGTGGGGGTTCTTGGAATCACTGGCTTTTGCCGACTAT
 GGATCCCTGGATAGAGTGAGATAGCCAGGCCCTCTATTACTCTTATCAA
 GGAGACCCAGAGGGCAAATACTTATCTAGTATTATGGGNACCAGAGGCAG
 AAAAGCCTTCCAAACCTTAAAGGAGACCCTAGTACAAGCTCCAGCTTTAA
 GCCTTCCCAACAGGACAAANCTTCTCTTTATATGTACAGAGAGAGCAGGAA
 TAGCTCCTGGAGTCCTTACTCAGACTTTTGACGACCCACGGCCAGTGGC
 RTACCTAAGTAAGGAAATTGATGTAGTAGCAAAAGGCTGGCCTCACTGTTT
 ATGGGTAGTTGCGGCTGTGGCAGTCTTACTGTCAAAGGCTATCAAAATAAT
 ACAAGGAAAGGATTTCACTATCTGGACTACTCATGAGGAAAATGGCATATT
 AGGTGCCAAAGGAAGTTTTTGGCTATCAGACAACCACCTGCTCAGATTCCA
 GGCACTACTGATTGAGAGACCAGTGCTTTAAATATGTATGTGTGTGTGG
 CCCTCAACCCTGCCACTGTTCTCCAGAAGATGGAGAACCAATGAAGCATT
 ACTGTCAACAAATTAGAGTCCAGAGTTATGCTGCCTGAGAGGATCTCTTAG
 AAGTCCCCTTAGCTAATCCTGACCTTAACCTATATGCTGATGGAAGTTCAC
 TTGTGGAGAATGGGATACGAAAAGCACATTATGCCATAGTTAGTGAGGTA
 ACAGTACTTGAAAGTAAGCCTATTCCCCCATGGACCAGAGCCCAGTTAGCA
 GAACTAGTGGCACTTACCCAAGCCTTAGAACTAGGAAAAGGGAAAAATAAT
 AAATGTGTATACAGATAGCAAGTATGCTTATCTAATCCTACATGCCCATGC
 TGCAGTATGGAAAGAAAGGGAGTTCTAACCTCTGGGGGAACCCCCATTA
 AATACCACAAGGCAAATCATGGAGTTATTGCATGTAGTGCAAAACCTCAA
 GTAGGTGGCAGTTTTACTGCCTGAAGCTATGGGGAAGGAGAGAGGAGA
 ACAGCAGCATAAGTGGCTAGCAGAGGCAGCGAAAGACTAGCAGAGAGGA
 GAGGTAGGGGAAAGACAGAAAGTCAAAGAAAAGAAGTCAAAGACAGACA
 GAGAAAGAGACAGAGGGAGCCAGAGAGAAAAGAAAAGAGAGAACGAAAGA
 GACAGAATGTCAAAGAACAGAAGAGAGAGGCGAGCGCCAGAAGAGTTAAG
 AAAGTGAGAAAGAGAGATGGAAATAGTAAAGAAAAAACAGTGTACCCTAT
 TCCTTTAAAAGCCAGGGTAAATTTAAAACGTATAATTTTATAATTGGAAGG
 TCTTCTCCATAACCTATAACATTAAAATACCACCTTGTTGTCAGTGTAAC
 AAGAGCATAGCCCCAAAAGCACTGAGGCCACTGACAACCCATAGCCTTCCT
 ATCAAAAATCCTTAACCTCTGCAGGTTTCCTAACAGGGGATCTAAATCTCAA
 CTAATCACCATACAATGGTCCGACCAGACCTAGGAGCGACTCCCCTCAGG
 ACAGAAGGATGGATGGTTCTCCAGGCCATTAAGGGAAAGAGACACAAT
 GGGTATTAGTAAGTGATAAGGGAACTCTTGTAAGAAGCAGTTAGGAAGATT
 GCCTAATATTTGGTCTGCTCAAATGTGCCAGCTGTTTGCACTCAGCTAAAC
 CTAAATTAATTACAGAATTAGGAAGGAGCCATCTATACCAATTCTGAGTT
 AATATGAGCTGAACAAGTTCTTATTAATAGCAAAGAATCATTGAAATCTCA
 AACTTGCAAAGTTTTCAACAAAAGTAAAGTTTGCTGAAAGTTAGCAGTGTA
 ACATGTATTATCCTAACTTCTAATCTTGTTGGAATCAGACCCTATCAGTGC
 CCCTCAAAGCTGAAGTCCATCAGCATATGGCCATACAATAATACCCTAT
 TTATAGGGTTAGGAATGGCCACTGCTACAGGAATGGGAGTAACAGGTTTAT
 CTACTTCATTATCCTATTACCACACACTCTTAAAGGATTTCTCAGACAGTTT
 ACAAGAAATAACAAAATCTATCCTTACTCTNTARTCCCAAATAGRTTCTTT
 GGCAGCAGTGACTCTC

FIG. 17

FIG. 18

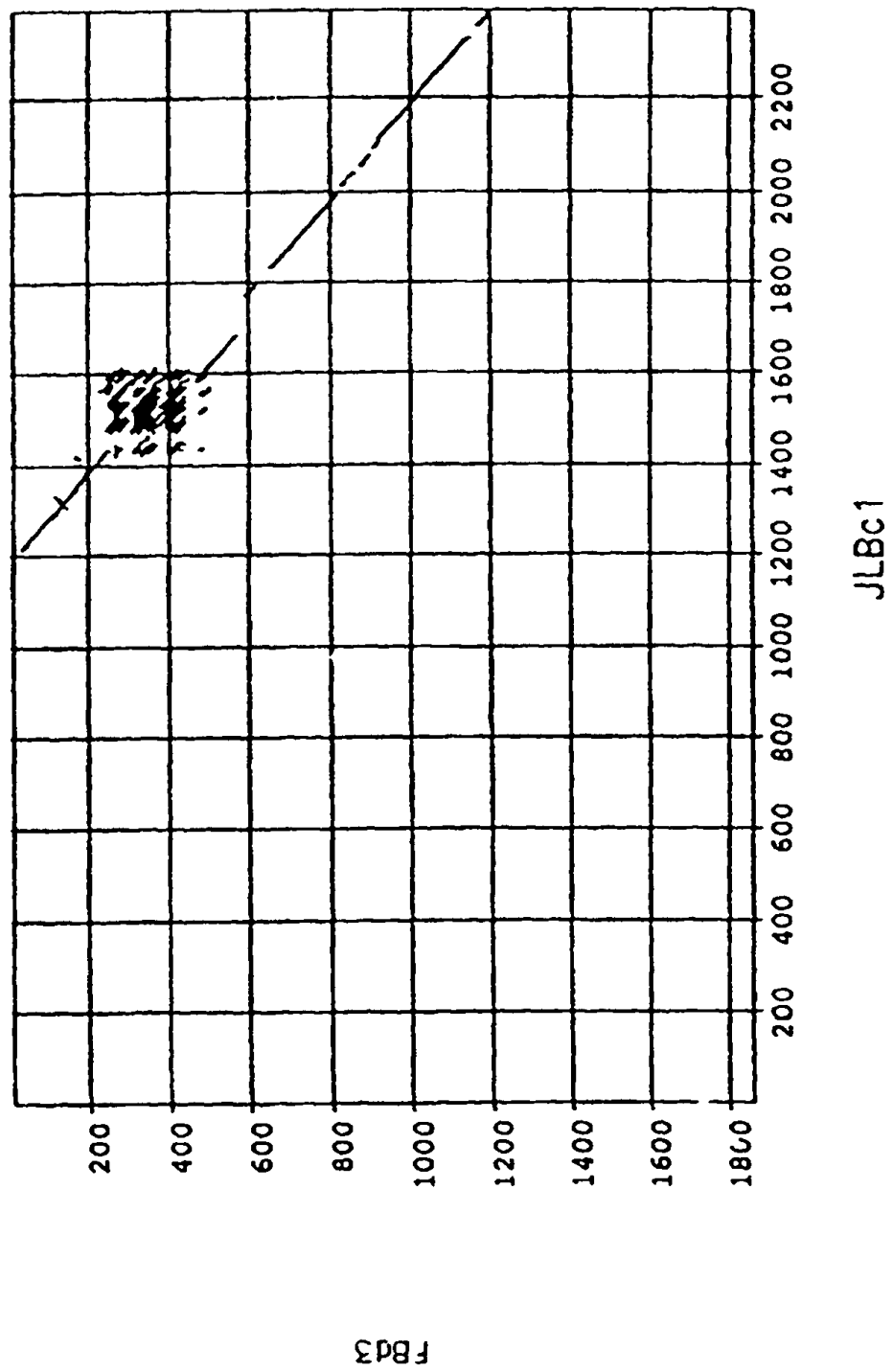
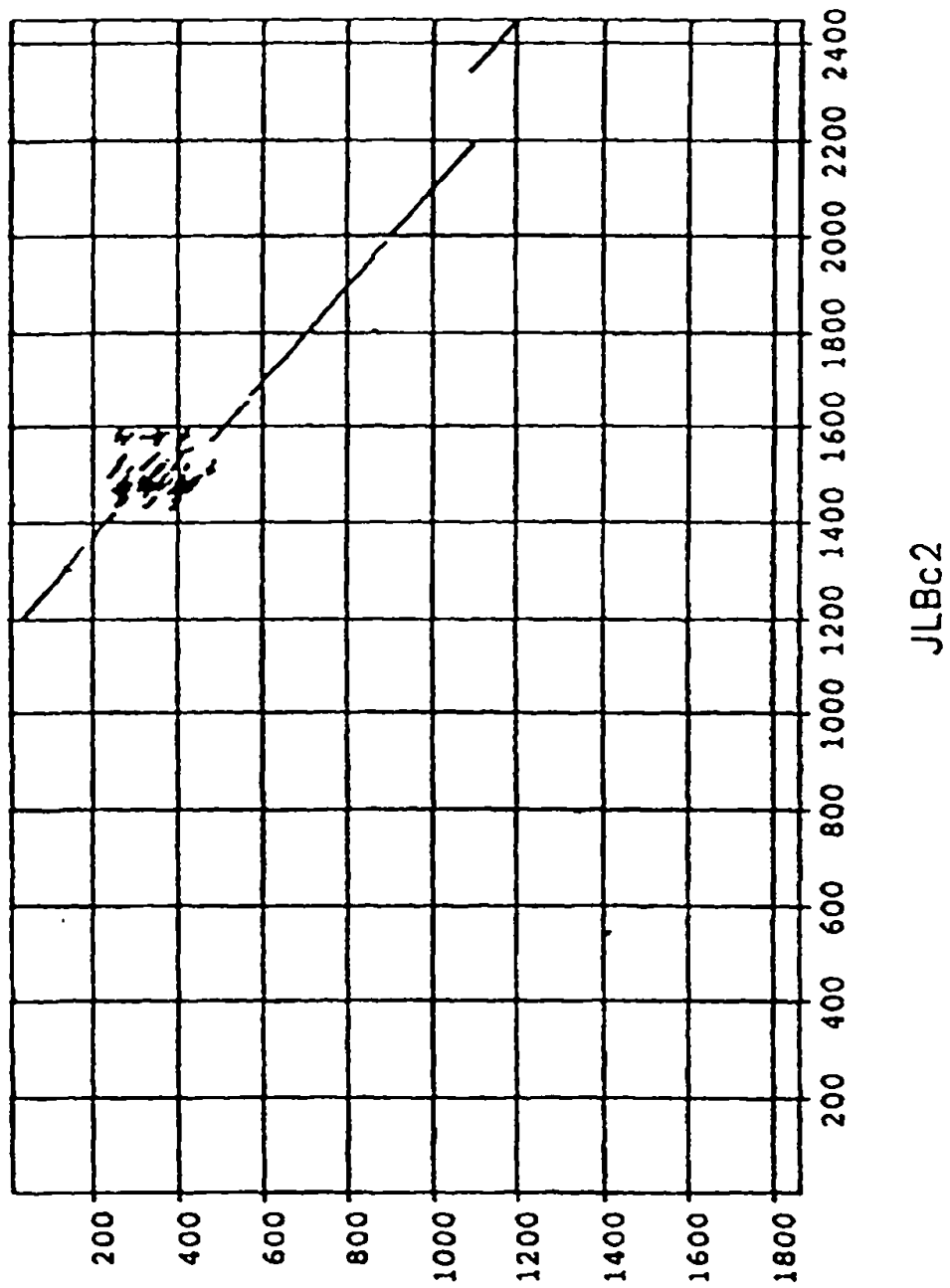


FIG. 19



FBd3

FIG. 20

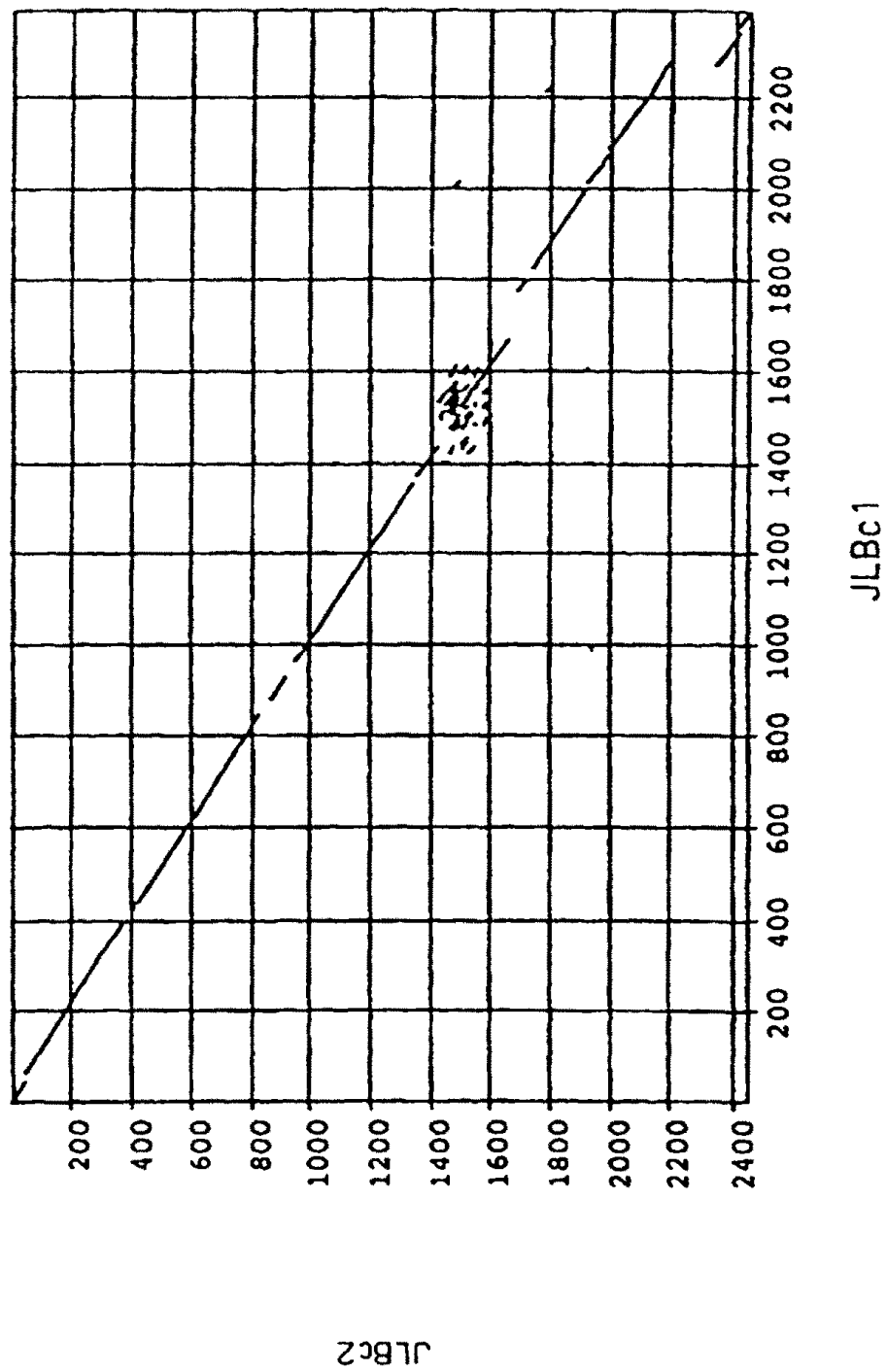
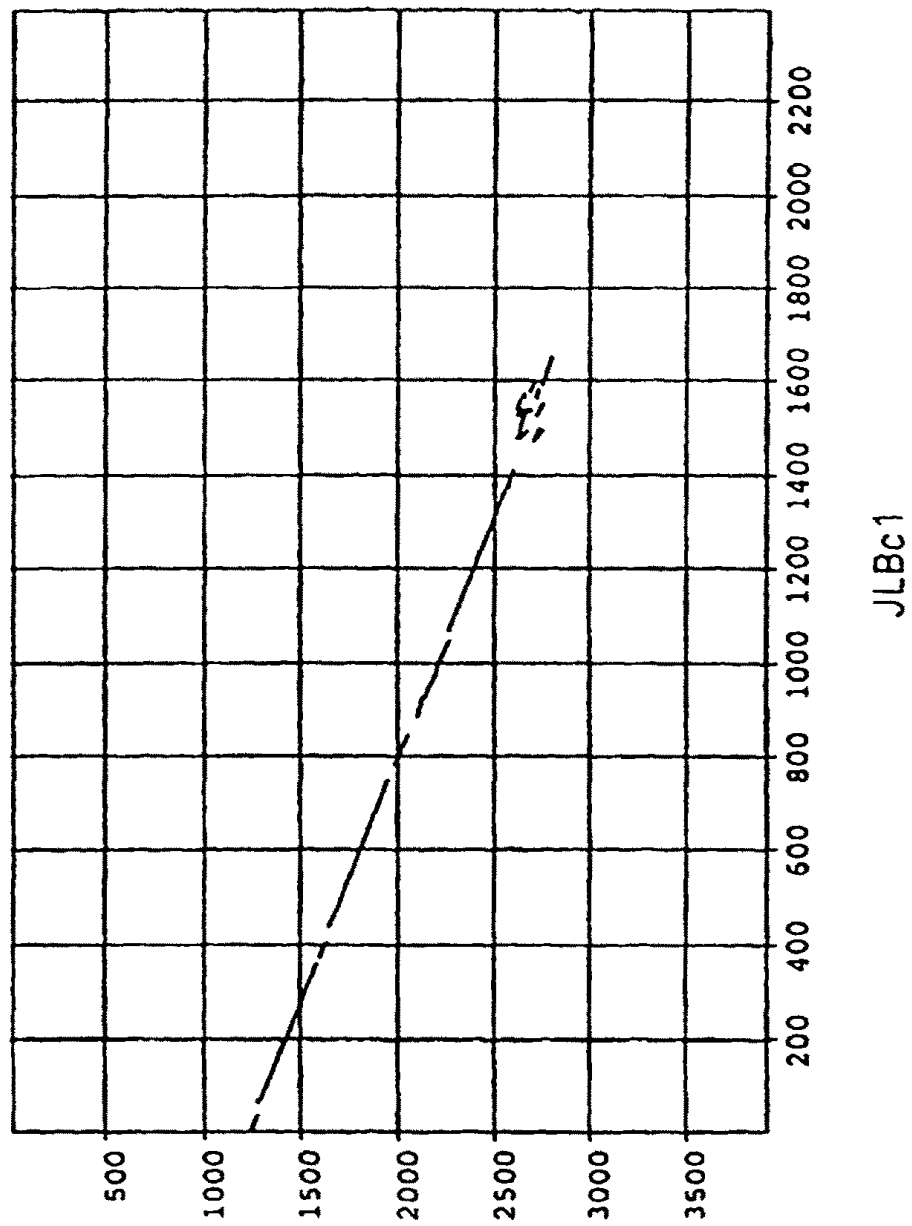
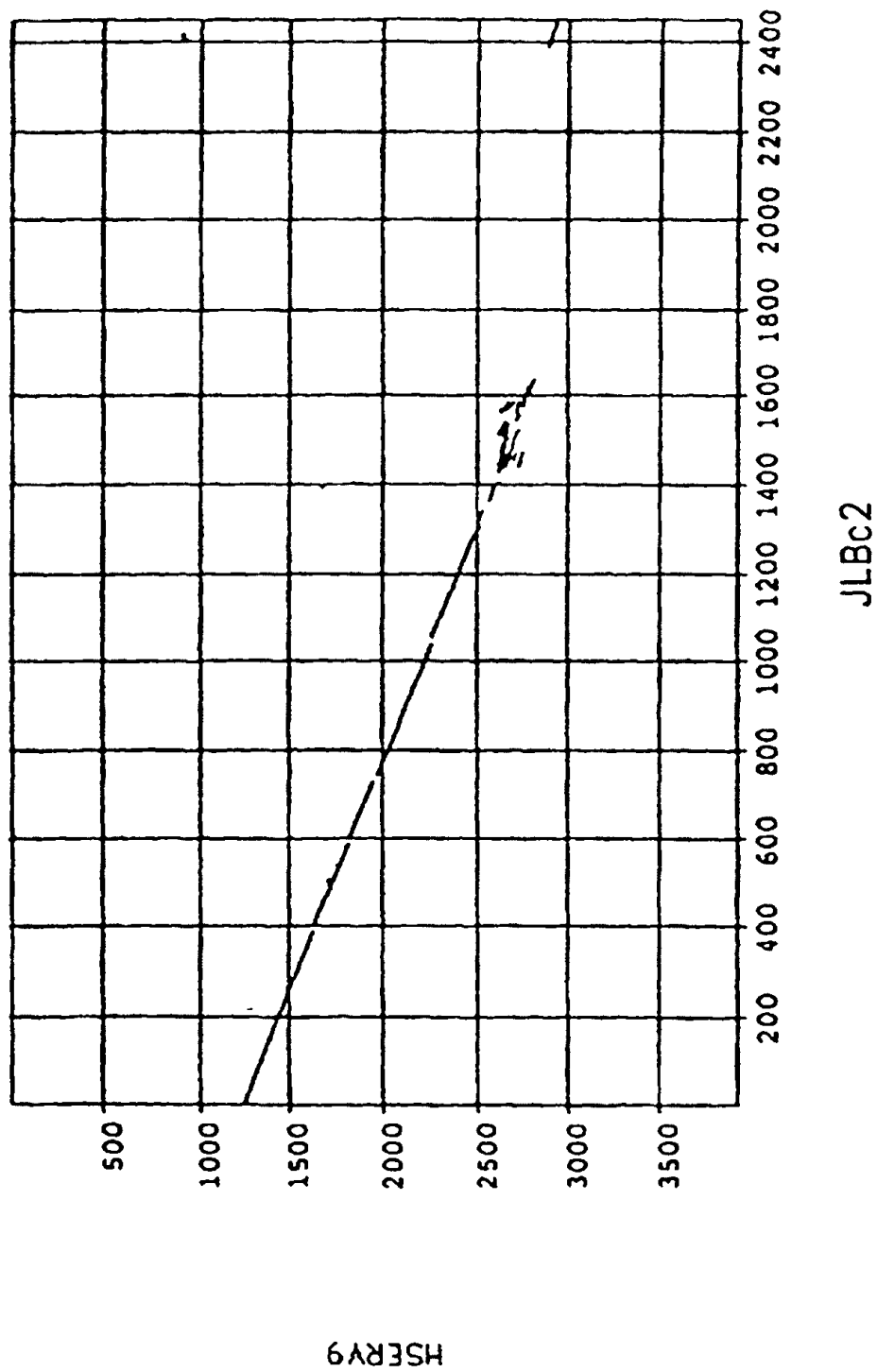


FIG. 21



HSERV9

FIG. 22



```

1  TTCCTGAGTT CTTGCACTAA CCTCAAATGA GAGAAGTGCC CCCATAACTG CAAACCCAAGA
61  GTTTGGCGAT CCCTGGTATC TCAGTCAGGT CAATGACAGG ATGACAACAG AGGAAAGATA
121 ATGATTCCCC ACAGGCCAGC AGGCAGTTCC CAGTGTAGAC CCTCATTAGG ACACAGAATC
181 AGAACATGGA GATTGGTGCC GCAGACATTT GCTAACTTGC GTGCTAGAAG GACTAAGGAA
241 AACTAGGAAG ATATGAATTA TTCAATGATG TCCACTATAA CACAGGGGAA AGGAAGAAAA
301 TCCTACTGCC TTTCTGGAGA GACTAAGGGA GGCATTGAGG AAGCATACCA GGCAAGTGGA
361 CATTGGAGGC TCTGAAAAAG GGAAAAGTTG GGAAAAGTAT ATGTCTAATA GGGCTTGCTT
421 CCAGTGTGGT CTACAAGGAC ACTTTAAAAA AGATTGTCCA ATAGAAATAA GCCACCACCT
481 CGTCCATGCC CCTTATGTCA AGGGAATCAC TGGAAGGCCC ACTGCCCCAG GGGATGAAGG
541 TCCTCTGAGT CAGAAGCCAC TAACCAGATG ATCCAGCAGC AGGACTGAGG GTGCCCCGGG
601 CAAGCGCCAG CCCATGCCAT CACCCTCACA GAGCCCCAGG TATGCTTGAC CATTGAGGGT
661 CAGAAGGGTA CTGTCTCCTG GACACTGGCG GGCCTTCTCA GTCTTACTTT CCTGTCTTGG
721 ACAACTGTCC TCCAGATCTG TCACTGTCCG AGGGGTCCCTA GGACAGCCAG TCACTAGATA
781 CTTCTCCCAG CCACTAAGTT GTGACTGGGG AACTTTACTC TTCCACATGC TTTTCTAATT
841 ATGCCTGAAA GCCCCACTCT CTTGTTAGGG GAGAGACATT CTAGCAAAAG CAGGGGCCAT
901 TATACATGTG AATATAGGAG AAGGAACAAC TGTTTGTTGT CCCCTGCTTG AGGAAGGAAT
961 TAATCCTGAA GTCCGGGCAA CAGAAGGACA ATATGGACAA GCAAAGAATG CCCGTCCTGT
1021 TCAAGTTAAA CTAAAGGATT CCACCTCCTT TCCCTACCAA AGGCAGTACC CCCTCAGACC
1081 CGAGACCCAA CAAGAACTCC AAAAGATTGT AAACGACCTA AAAGCCCAAG GCCTAGTAAA
1141 ACCAAGCAAT AGCCCTTGCA AGACTCCAAT TTTAGGAGTA AGGAAACCCA ACGGAC

```

SEC ID n° 56 (GM3)

FIG. 23

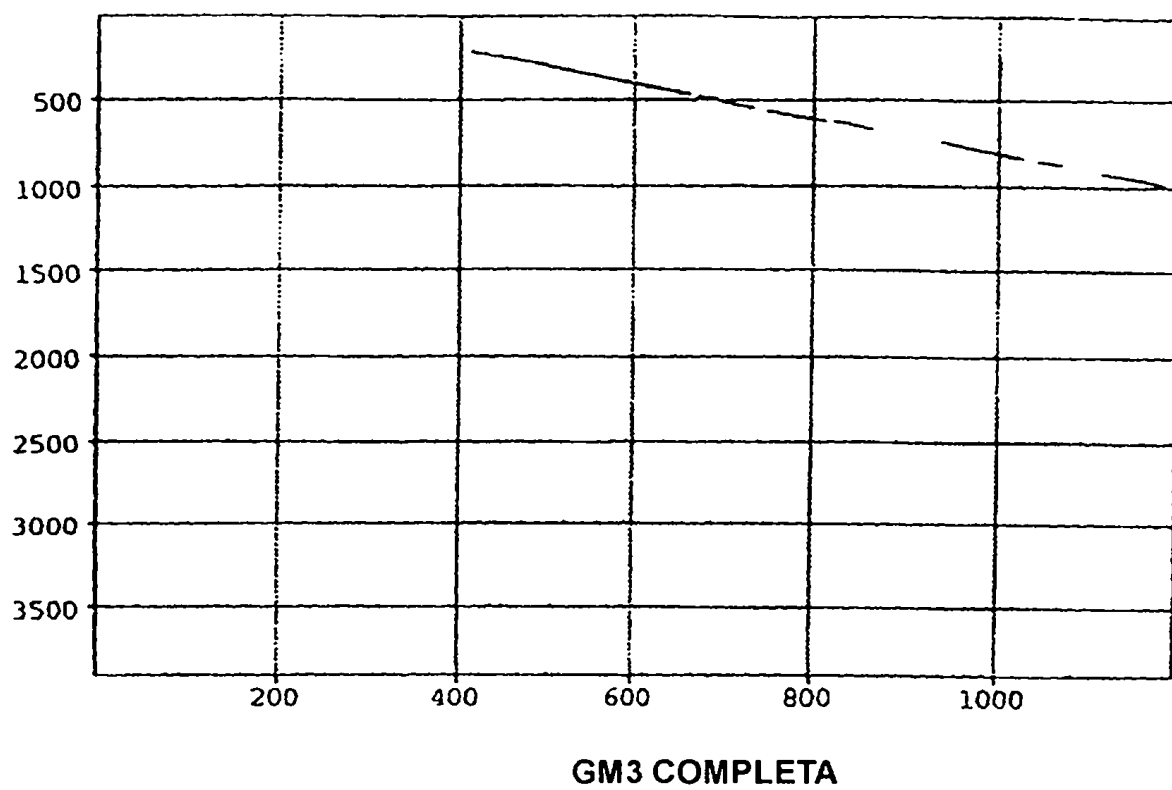
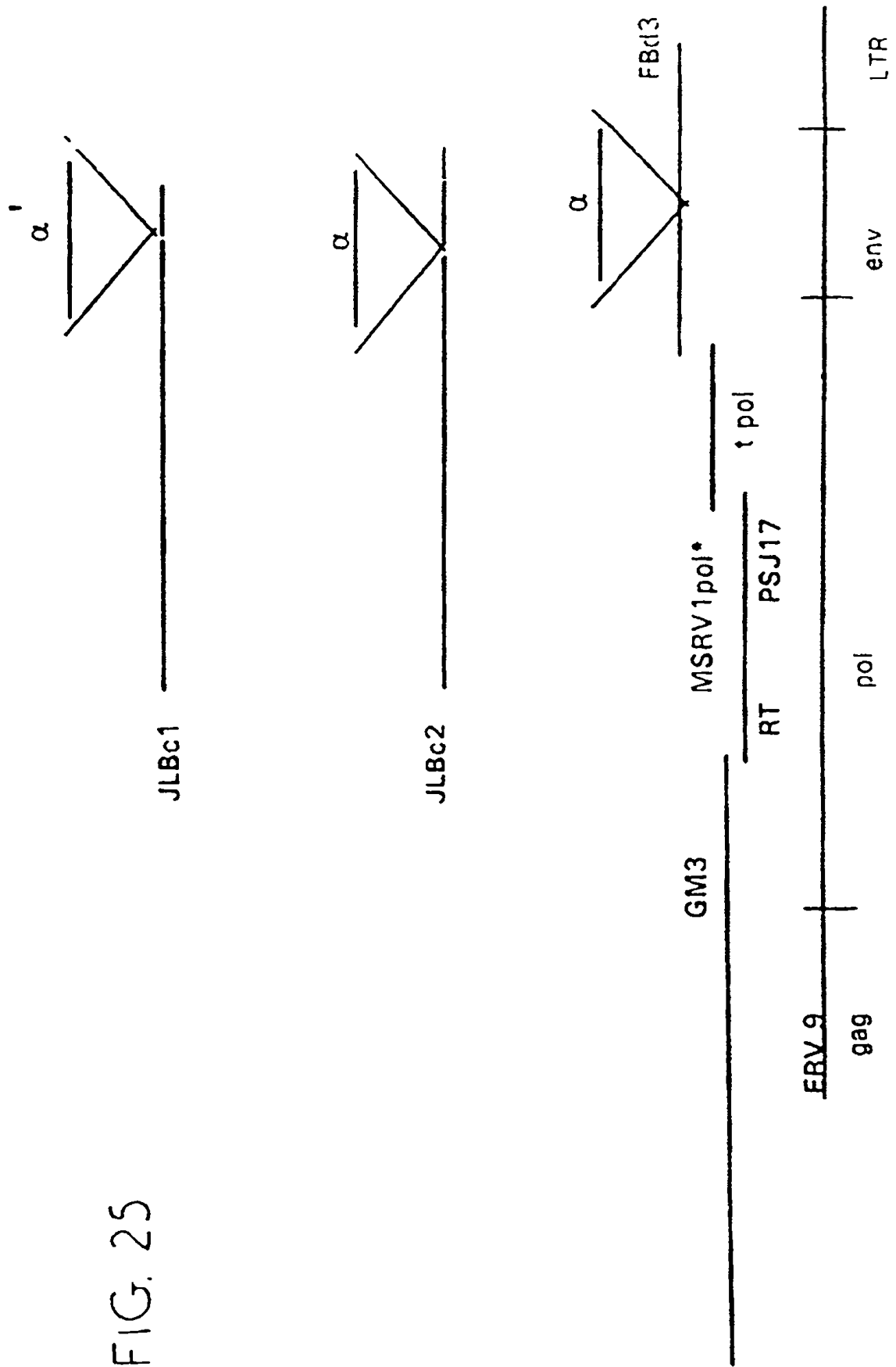


FIG. 24

FIG. 25



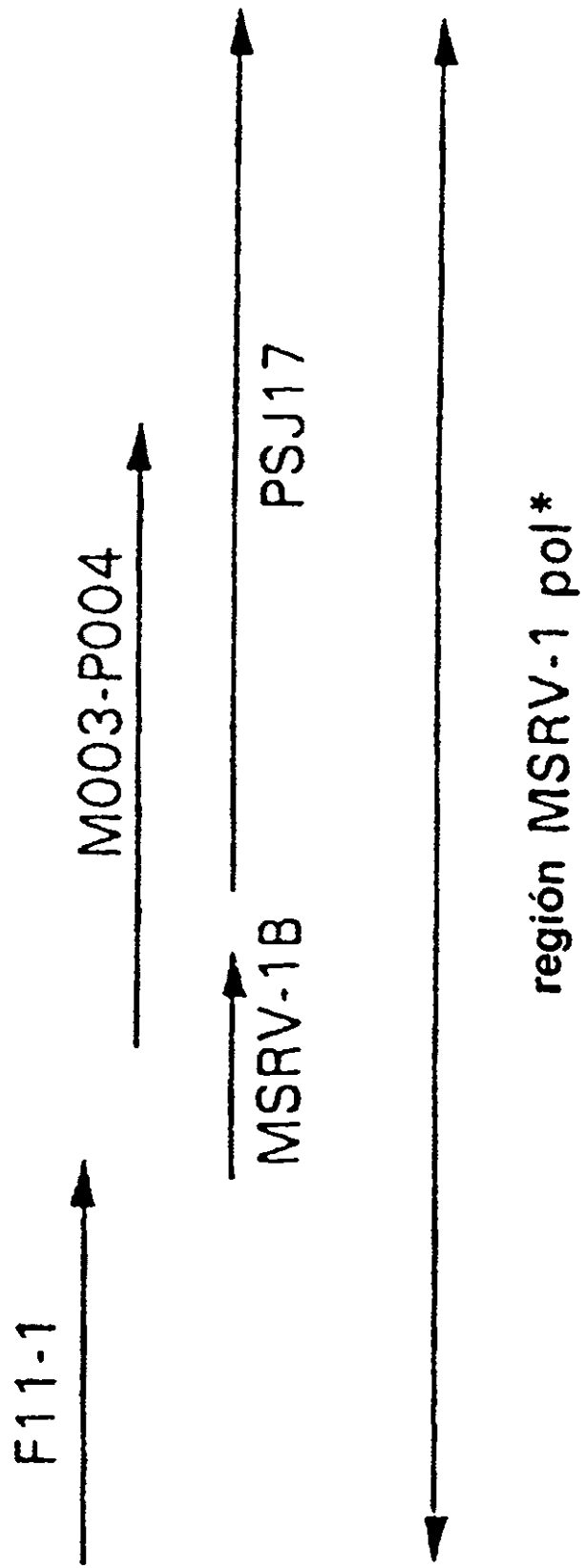


FIG. 26

FIG. 27a

SEC ID n° 57 (POL)

ATG	ATC	CAG	CAG	CAG	GAC	GAC	NGA	GAG	TGC	GGG	GGA	GGG	GGA	GGG	CCA	GCC	CAT	GCC	ATC	ACC	CTC	ACA	GAG	CCC	CAG	GTA	TGC	TTG	ACC	ATT	GAG	90		
M	I	Q	Q	Q	D	Q	X	G	C	P	G	Q	A	G	A	P	A	H	A	I	T	L	T	E	P	Q	V	C	L	T	I	E		
GGT	CAG	AAG	GCT	NAC	TGT	CTC	CTC	CTG	GAC	ACT	GGC	GGN	GGC	GGC	TTT	TTA	GTC	TCA	CTT	TGC	TGT	GGT	GGC	ACT	TTA	CTC	TTT	AGA	TCT	GTC	ACT	180		
G	Q	K	G	X	C	L	L	L	D	T	G	G	A	F	S	V	L	L	L	S	C	P	G	Q	L	S	S	S	R	S	V	T		
GTC	CCA	GGG	GTC	CTA	CCA	CAG	CCA	CTC	ACT	ACT	AGA	TAC	TTT	TTT	F	S	Q	P	L	S	C	D	W	G	T	L	L	CTC	TTT	CCA	CAT	TTT	270	
V	R	G	V	L	G	Q	Q	P	V	T	R	Y	F	F	S	Q	P	L	L	AGT	TGT	GAC	TGG	GGG	ACT	TTA	CTC	TTT	CTC	CAI	CAI	TTT	F	
CTA	ATT	ATG	ACT	GAA	ACC	ACT	CTC	TTG	TTG	TTG	GGG	AGA	GAC	ATT	CTA	CCA	AAA	CCA	GGG	GGC	GGC	GGG	GGC	ACT	ATA	CAI	GGT	ATG	ATT	ATA	GGA	GGA	360	
L	I	M	F	E	S	P	T	L	L	L	L	G	R	D	I	L	A	K	A	G	A	I	I	H	V	N	I	C	E	G	E	G		
ACA	ACT	GTT	TGT	TGT	CCC	CTG	CTT	GAG	GAA	CCA	CCA	GGG	CCA	ACA	GAA	GGG	CCA	ACA	ACA	TAT	GGA	CCA	CCA	TAT	GGA	CCA	CCA	ACA	AGA	ATT	GGC	GGC	450	
T	T	V	C	C	P	L	L	E	E	G	I	N	P	E	V	R	A	T	E	G	Q	Y	G	Q	A	Q	A	K	N	A	R	R		
QCT	GTA	QVQ	KL	AAK	CTA	AGK	GAT	TCC	ACC	TCC	TTT	CCC	TAC	CCA	AGG	ONG	TAC	CCC	CTC	AGA	CCC	GAG	ACC	CCA	CCA	CCA	CCA	CCA	CCA	CTC	CCA	AAG	540	
P	V	Q	V	K	L	K	D	S	T	S	F	P	Y	Q	R	Q	Y	P	L	R	P	E	T	Q	Q	Q	Q	E	L	Q	K	K		
AIT	ATA	ASG	CAC	CTA	AAA	GGC	CAA	GGC	CTA	GTA	AAA	CCA	ACC	ANT	ACC	CTT	TCC	AGS	ACT	CCA	ATT	TTA	GGA	GTA	GGG	AAA	CCC	ANC	GGG	GGG	GGG	630		
I	V	K	D	L	K	A	Q	G	L	V	K	P	S	N	S	P	C	K	T	P	I	L	G	V	R	K	P	N	G	N	G	G		
CAG	TGG	AGG	ATA	CTG	CAA	CTC	AGG	ATT	ATC	ANT	GAT	GCT	GTT	GTT	V	V	P	L	Y	P	A	V	P	A	V	P	N	P	Y	T	V	L	S	720
Q	W	R	L	V	Q	E	L	R	I	I	N	E	A	V	V	P	L	Y	P	A	P	A	G	CTA	GGT	CTT	GGT	CTT	GGT	CTT	GGT	CTT	GGT	
CAA	ATA	CCA	CAG	GAA	CCA	CAG	TGG	TTT	ACA	GTC	CTG	GAC	CTT	AGG	CAI	GGC	TTT	TTT	TTT	CTC	ATC	CTT	GTA	GGT	CTT	GGT	CTT	GGT	CTT	GGT	CTT	GGT	810	
Q	I	P	E	F	A	E	W	F	T	V	L	D	L	K	D	A	F	F	C	I	P	V	R	P	D	S	Q	F	L	F	L	L		
TTT	GGC	TTT	GAA	GAT	ACT	TTG	ANC	CCA	ACC	TCT	CCA	CTC	CCA	CTC	TTG	TTT	TTT	TTT	TTT	TTT	TTT	TTT	TTT	TTT	TTT	TTT	TTT	TTT	TTT	TTT	TTT	TTT	900	
F	A	F	E	D	P	L	H	P	T	S	Q	L	T	N	T	V	L	P	Q	P	Q	G	F	R	D	S	P	H	L	F	G	G		
CAG	CCA	TTA	GGC	CCA	GAC	TTG	AGT	CCA	TTT	CTA	TAC	CTG	GAC	ACT	CTT	GTC	CTT	CAG	TAC</															

FIG. 27b

SEC ID n° 57 (POL)

TCA	CAG	GAG	ATT	AGA	TAC	TAA	GCG	CTA	AAA	TTA	TCC	AAA	GEC	AOC	REG	GCC	CTC	AGT	GAG	GAA	GGT	ATC	CAG	CCT	ATA	CTG	GCT	TAT	OCT	1170	
S	Q	E	I	R	Y	X	G	L	K	L	S	K	G	T	R	A	L	S	E	E	E	I	Q	P	I	L	A	Y	P		
CAT	COE	APA	ACC	CTA	AAG	CAA	CTA	AGA	GCG	TTC	CTT	GGC	ATA	ACA	GGT	TTC	TCC	CGA	AAA	CAG	ATT	CCC	AGS	TAC	ASC	CCA	ATA	CCC	AGA	1260	
H	P	K	T	L	K	Q	L	R	G	F	L	G	I	T	G	F	C	R	K	Q	I	P	R	Y	X	P	I	A	R		
CCA	TTA	TAT	AGA	CTA	ATT	ANG	GAA	ACT	CAG	AAA	GCC	AAT	AOC	TAT	TTA	GTA	AGR	TGG	ACA	GGT	NCA	GAG	TTC	TTC	CAG	CCC	CTA	AAG	1350		
F	L	Y	C	L	I	X	E	T	Q	K	A	N	T	Y	L	V	R	W	T	P	T	E	V	A	F	Q	A	L	K		
ANG	GCC	CTA	ACC	CAA	GCC	CCA	GTC	CTC	ACC	TTG	ACC	ACA	GEG	CAA	GAT	TTT	TCT	TTA	TAT	TAT	CCC	ACA	GAA	AAA	ACA	ATA	GCT	CTA	1440		
K	A	L	T	Q	A	P	V	F	S	L	P	T	G	Q	D	F	S	L	Y	A	T	E	K	T	G	I	A	L	G		
GTC	CTI	ADG	CAG	STC	TCA	GGG	ATG	ACC	TTG	CAA	CCC	GTG	GTA	TAC	CTG	AGT	ANG	GAA	ATT	GAT	GTA	GTG	GCA	AAG	GGT	TGG	CCG	CAT	NGT	1530	
V	L	T	Q	V	S	G	M	S	L	Q	P	V	V	Y	L	S	K	E	I	D	V	V	A	K	G	W	P	H	X		
ITA	TGG	GTA	ATG	GAG	CCA	GTA	GTC	TAA	GTA	TCT	GAA	GTA	GGA	GTT	AAA	ATA	ATA	CAG	ACA	GAT	CTT	NCT	GTG	TGG	TGG	ACA	TCT	CAT	1620		
L	W	V	H	X	A	V	A	V	X	V	S	E	A	V	K	I	I	Q	G	R	D	L	X	V	W	T	S	H	D		
GTG	PAC	GGC	ATA	CTC	ACT	ACT	AAA	GGA	GSC	TTG	TGG	L	S	D	N	H	L	L	X	Y	Q	A	L	L	L	L	E	E	P	1710	
V	N	G	I	L	T	A	K	G	D	L	W	L	S	D	N	H	L	L	X	Y	Q	A	L	L	L	L	L	E	E	P	V
CTG	NCA	CTG	CCC	ACT	TGT	CCA	ACT	CTT	AAA	CCC	AAA	CTT	ATG	CTG	CCC	ACA	AGS	ATC	TTT	MTA	GAG	GTG	CCC	TTA	CCC	AAC	CCG	GNC	CTC	1800	
L	X	L	R	C	A	T	L	K	P	K	L	L	H	L	P	R	R	I	F	X	E	V	P	L	A	N	P	D	L		
AMC	TAT	ATA	TAT	ACT	GAT	GGA	AGT	TGG	TTT	GTA	GAA	ANG	GGA	CTA	CAA	AGG	GNA	GGA	DAT	NCC	ATA	GGT	GGT	AGT	CAT	ANA	GCA	GTA	CTT	1890	
Y	N	Y	I	Y	T	D	C	S	S	F	V	E	K	G	L	Q	R	X	G	Y	X	I	G	V	S	D	K	A	V	L	
GAA	AGT	ASG	CCC	CTT	CCC	CAG	GGA	CCA	GGG	CCC	CCA	GGP	L	A	E	L	V	A	L	L	T	P	R	A	L	E	L	W	K	G	1980
E	S	K	P	L	P	Q	G	G	P	A	P	P	L	A	E	L	V	A	L	L	T	P	R	A	L	E	L	W	K	G	
ACG	ACG	ATA	ANT	GTG	TAT	ACA	GAT	ADC	ANG	TAT	GGT	TAT	CTA	ATC	CCA	AAT	CCC	CAT	GGT	TAT	CTA	ATC	CGA	ANT	CCC	CAT	GGT	CCA	2070		
R	R	I	N	V	Y	T	D	S	K	Y	A	Y	L	I	R	N	A	H	V	V	Y	L	I	R	N	A	H	V	A		
ATA	TGG	AAA	GAA	AGG	GAG	TTC	CTA	ACC	TCT	GGG	CCA	ACC	CTC	ACC	ACT	AAA	TAC	CAC	ANG	TTA	ATC	GAG	TTA	TTG	CAC						

FIG. 27c

SEC ID n° 57 (POL)

CTC	ANG	GAG	GAG	GTC	TTA	CAC	TCC	CAA	ACC	CAT	CAG	AAA	ACG	GAA	AGG	GGA	GGC	GAA	GGA	GAA	GGA	GGA	GGA	GGA	ANG	2250
L	K	E	V	E	V	L	H	C	Q	S	H	Q	K	R	E	R	G	E	E	E	Q	C	Q	Q	G	K
ACT	ACC	AGA	ANG	GAA	GAG	AAA	GAG	ACA	GAA	AGT	CAG	AGA	GAG	GAA	GAG	GAA	GAA	GAA	GAA	GAA	GAA	GAA	GAA	GAA	GAA	2250
T	S	R	K	E	R	E	K	E	T	E	S	Q	R	E	E	E	E	E	E	E	E	E	E	E	E	
AGA	CAG	AGA	UTC	ACA	GAG	ANG	GAA	AGA	GAG	AGA	AGA	GAC	AGA	GAA	TGA											
R	Q	R	V	R	E	K	E	R	E	R	G	R	D	K	E											
																									2391	

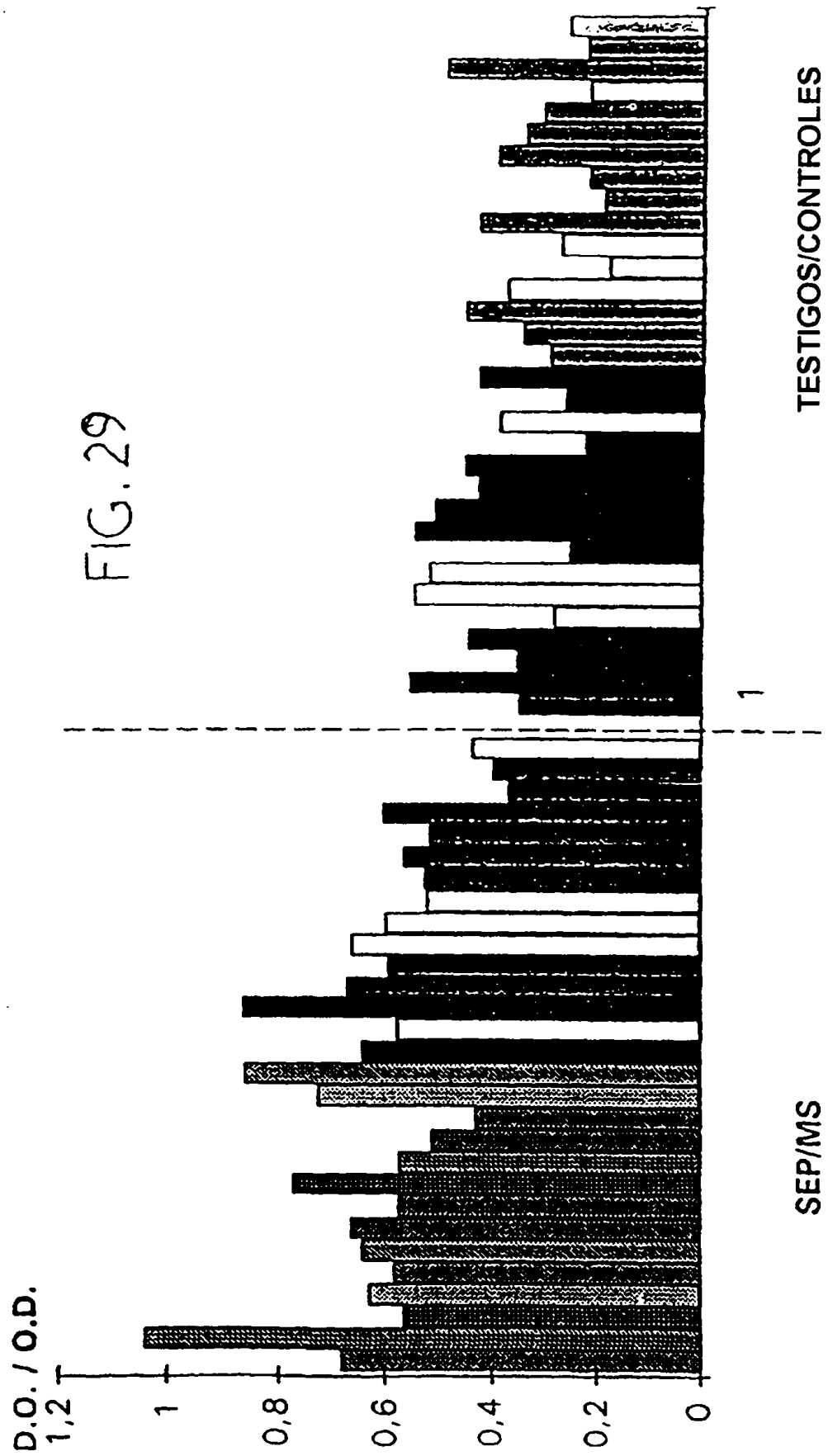
FIG. 28

GATGCCTTTTTCTGCATCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAG
ATCCTTTGAACCCAACGTCTCAACTCACCTGGACTGTTTTACCCAAGGGTTCAGGGA
TAGCCCCATCTATTTGGCCAGGCATTAGCCCAAGATGCCTTTTGCATCCCTGTACGTG
ACTCTCAATTCTTGTTTGCCTTTGCCTTTGAAGATGCTTTGAACCCAACGTCTCAACT
CACCTGGACTGTTTTACGCCAAGGGTTCAGGGATAGCCCCCATCTATTTGGC
CAGGCATTAGCCCAA

SEC ID n° 40

Asp-Ala-Phe-Phe-Cys-Ile-Pro-Val-Arg-Pro-Asp-Ser-Gln-Phe-
Leu-Phe-Ala-Phe-Glu-Asp-Pro-Leu-Asn-Pro-Thr-Ser-Gln-Leu-
Thr-Trp-Thr-Val-Leu-Pro-Gln-Gly-Phe-Arg-Asp-Ser-Pro-His-
Leu-Phe-Gly-Gln-Ala-Leu-Ala-Gln

SEC ID n° 39 (POL2B)



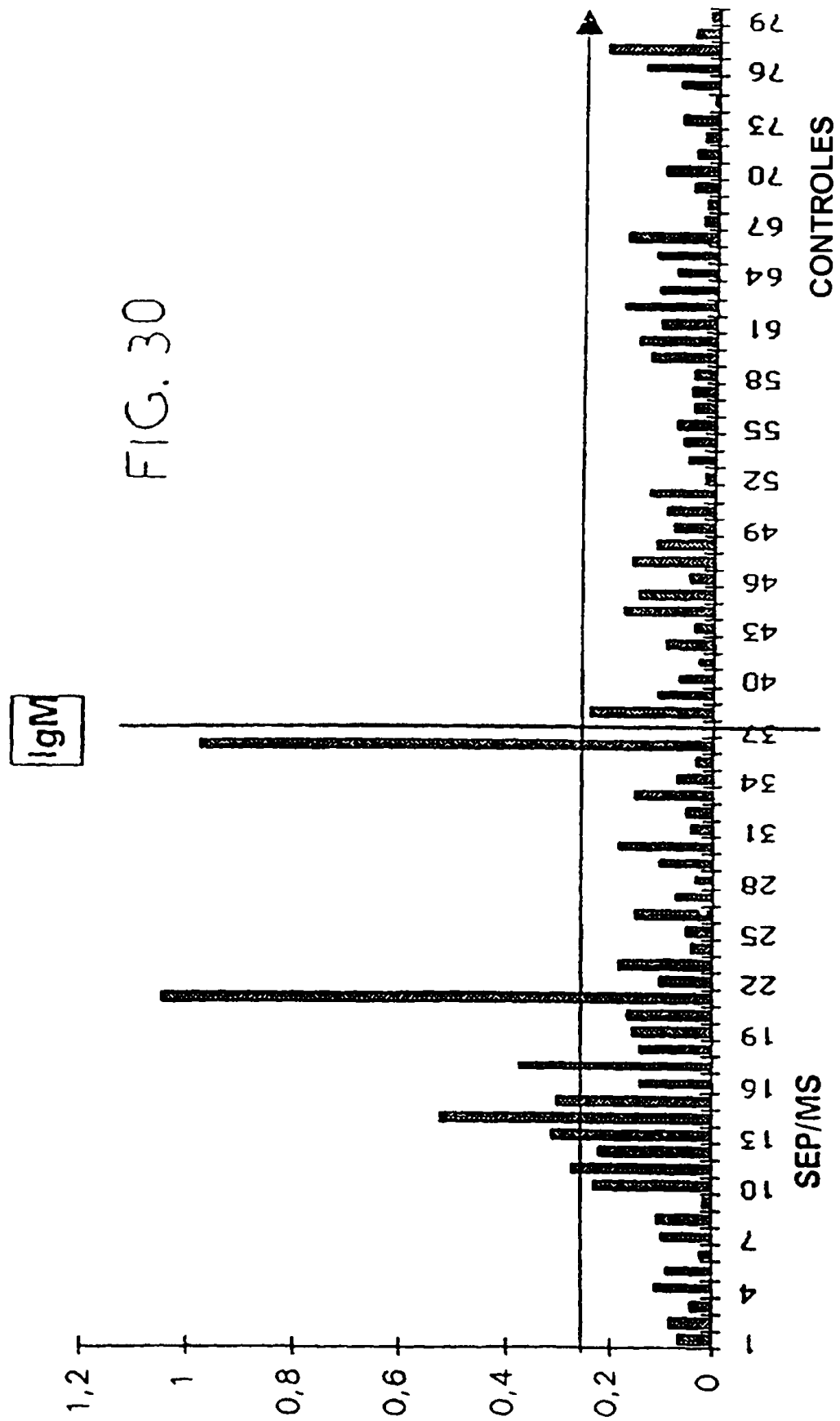


FIG. 31

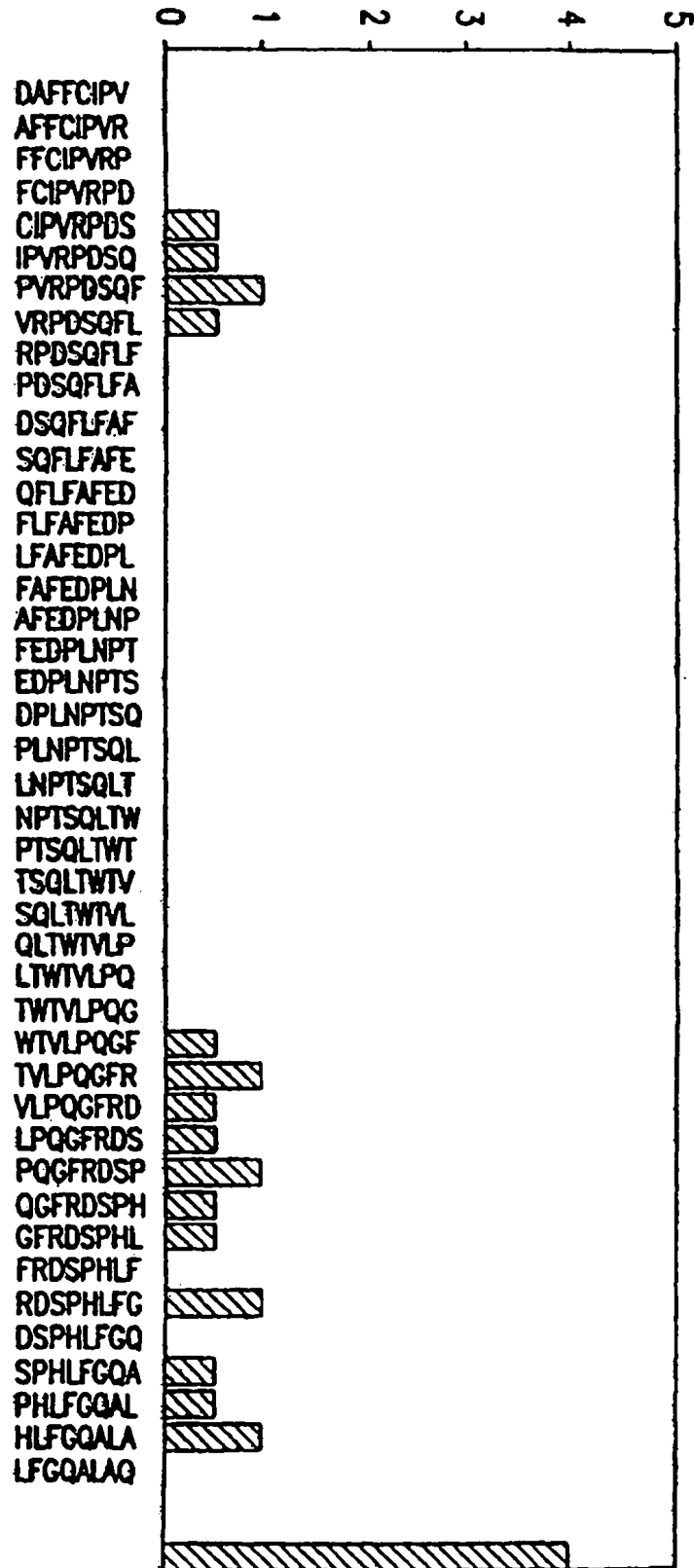
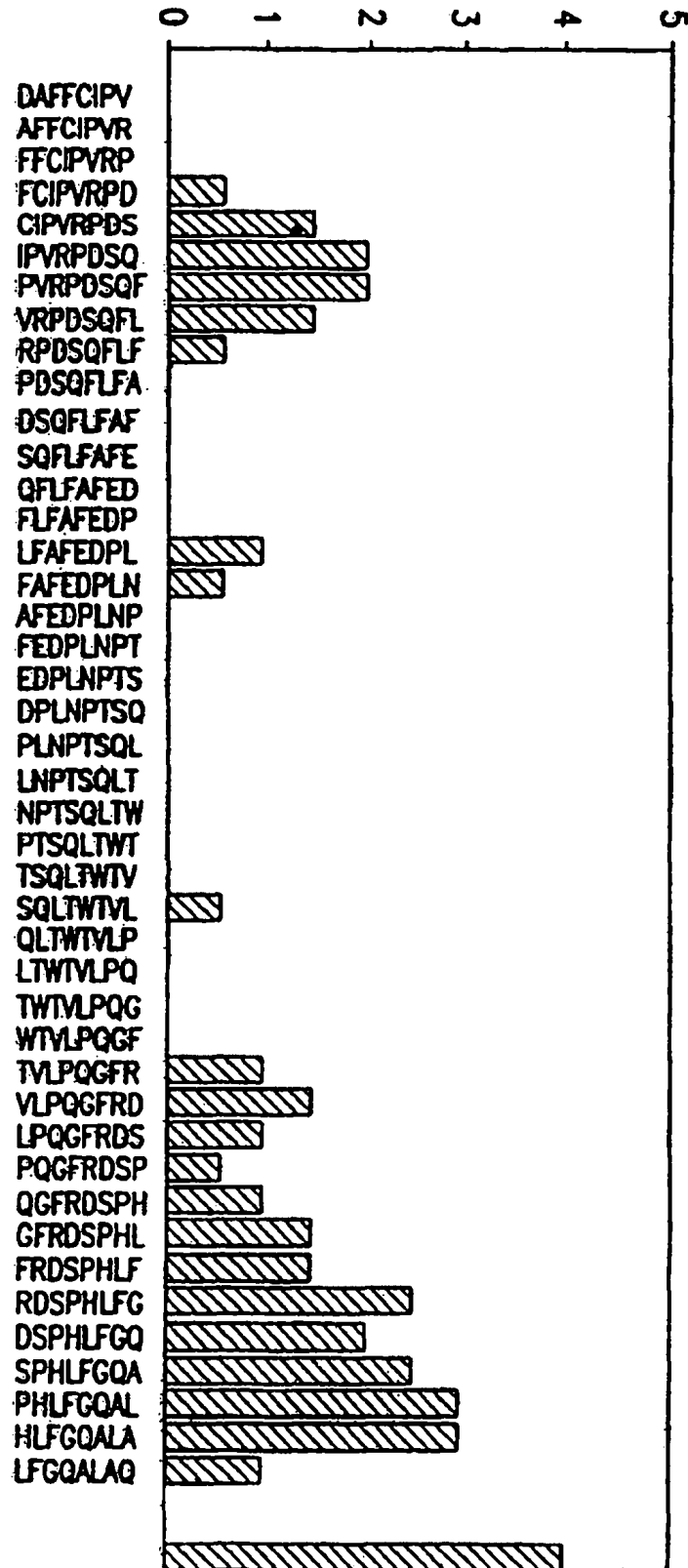


FIG. 32



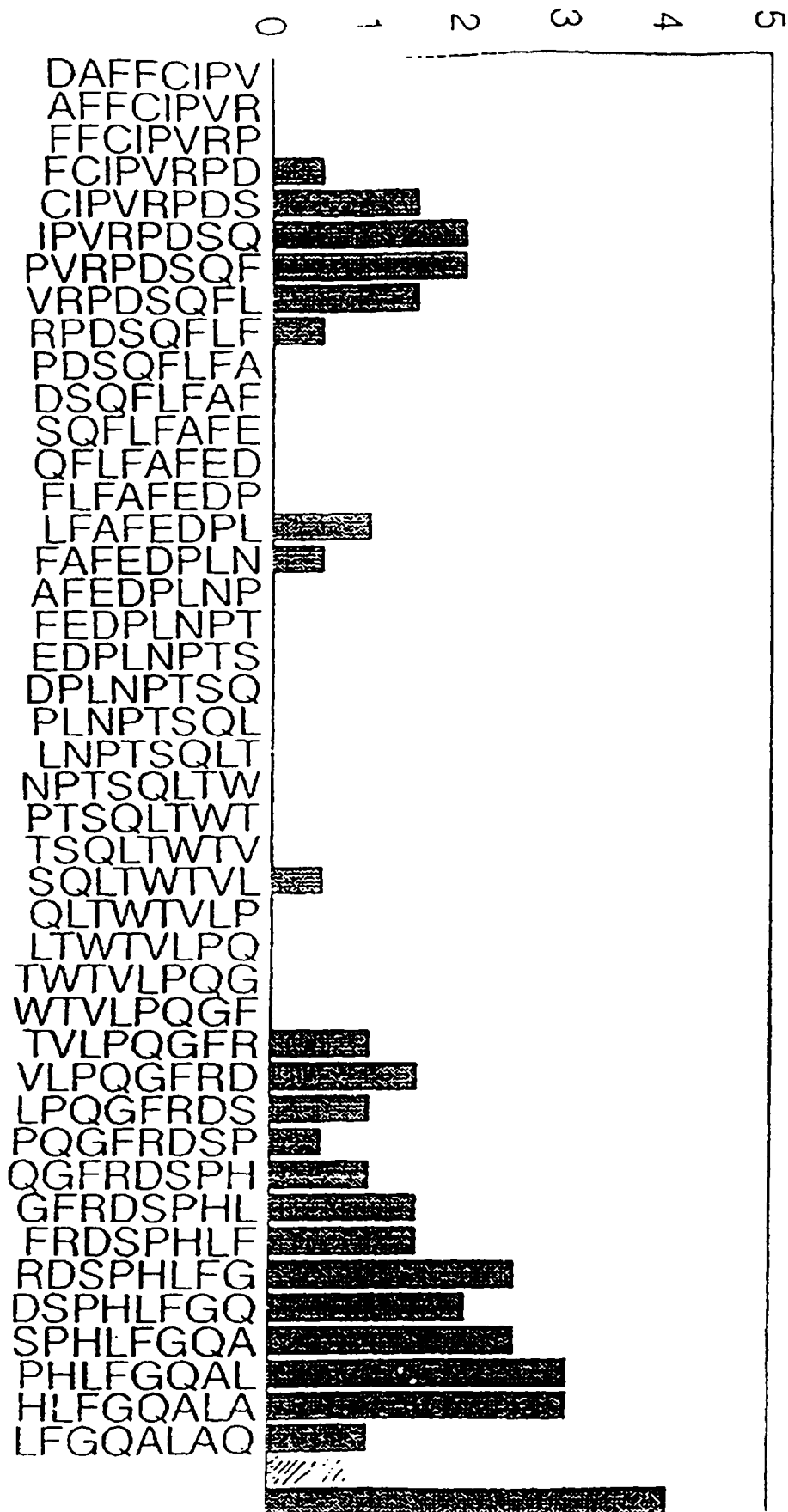


FIG. 33

FIG. 34

Cys-Ile-Pro-Val-Arg-Pro-Asp-Ser-Gln-Phe-Leu **SEC ID nº 41**

Val-Leu-Pro-Gln-Gly-Phe-Arg-Asp-Ser-Pro-His-Leu-Phe-Gly-
Gln-Ala-Leu-Ala **SEC ID nº 42**

Leu-Phe-Ala-Phe-Glu-Asp-Pro-Leu **SEC ID nº 43**
Phe-Ala-Phe-Glu-Asp-Pro-Leu-Asn **SEC ID nº 44**

FIG35

10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	
CTTCCCAAC	TAATAAGGAC	CCCCTTTCA	ACCCAAACAG	TCCAAAAGCA	50
L P Q L	I R T	P L S	T Q T V	Q K D	
F P N	. . G P	P F Q	P K Q	S K R T	
S P T	N K D	P P F N	P N S	P K G	
CATAGACAAA	GCAGTAAACA	ATGAACCAAA	GAGTGCCAAT	ATTCCCTGGT	100
I D K	G V N N	E P K	S A N	I P W L	
. T K E	. T M N	Q R V	P I F	P G	
H R Q R	S K Q	. T K	E C Q Y	S L V	
TATGCACCTT	CCAAGCGTG	GCAGAAGAAT	TGGCCCCAGC	CAGAGTGCAT	150
C T L	Q A V	G E E F	G P A	R V H	
Y A P S	K R W	E K N	S A Q P	E C M	
M H P	P S G G	R R I	R P S	Q S A C	
GTACCTTTTT	CTCTCTCACA	CTTGAAGCAA	ATTAAAATAG	ACNTAGGINA	200
V P F S	L S H	L K Q	I K I D	X G X	
Y L F	L S H T	. S K	L K .	T . V N	
T F F	S L T	L E A N	. N R	X R X	
ATTTCAGAT	AGCCCTGATG	GYTATATTGA	TGTTTTACAA	GGATTAGGAC	250
X S D	S P D G	Y I D	V L Q	G L G Q	
X Q I	A L M	X I L M	F Y K	D . D	
I X R .	P . W	L Y .	C F T R	I R T	
AATCCTTTGA	TCTGACATGG	AGAGATATAA	TATTACTGCT	AAATCAGACG	300
S F D	L T W	R D I I	L L L	N Q T	
N P L I	. H G	E I .	Y Y C .	I R R	
I L .	S D M E	R Y N	I T A	K S D A	
CTAACCTCAA	ATGAGAGAAG	TGCTGCCATA	ACTGGAGCCC	GAGAGTTTGG	350
L T S N	E R S	A A I	T G A R	E F G	
. P Q	M R E V	L P .	L E P	E S L A	
N L K	. E K	C C H N	W S P	R V W	
CAATCTCTGG	TATCTCAGTC	AGGTCAATGA	TAGGATGACA	ACCGAGGAAA	400
N L W	Y L S Q	V N D	R M T	T E E R	
I S G	I S V	R S M I	G . Q	R R K	
Q S L V	S Q S	G Q .	. D D N	G G K	
GAGAACGATT	CCCCACAGGG	CAGCAGGCAG	TTCCAGTGT	AGCTCCTCAT	450
E R F	P T G	Q Q A V	P S V	A P H	
E N D S	P Q G	S R Q	F P V .	L L I	
R T I	P H R A	A G S	S Q C	S S S L	
TGGGACACAG	AATCAGAACA	TGGAGATTGG	TGCCCCAGAC	ATTTA	495
W D T E	S E H	G D W	C R R H	L	
G T Q	N Q N M	E I G	A A D	I	
G H R	I R T	W R L V	P Q T	F	

FIG 36

	10	20	30	40	50	
	1234567890	1234567890	1234567890	1234567890	1234567890	
	CTTCCCCAAC	TAATAAGGAC	CCCCCTTTCA	ACCCAAACAG	TCCAAAAGGA	50
	L P Q L	I R T	P L S	T Q T V	Q K D	
	CATAGACAAA	GGAGTAACAA	ATGAACCAAA	GAGTGCCAAT	ATTCCCTGGT	100
	I D K	G V N N	E P K	S A N	I P W L	
	TATGCACTCT	CCAAGCGGTG	GGAGAAGGAT	TGGCCCCAGC	CAGAGTGCAT	150
	C T L	Q A V	G E E F	G P A	R V H	
	GTACCTTTTT	CTCTCTCACA	CTTGAAGCAA	ATTAAAATAG	ACCTAGGTAA	200
	V P F S	L S H	L K Q	I K I D	L G K	
	ATTCTCAGAT	AGCCCTGATG	GTTATATTGA	TGTTTACAA	GGATTAGGAC	250
	F S D	S P D G	Y I D	V L Q	G L G Q	
	AATCCTTTGA	TCTGACATGG	AGAGATATAA	TATTACTGCT	AAATCAGACG	300
	S F D	L T W	R D I I	L L L	N Q T	
	CTAACCTCAA	ATGAGAGAAG	TGCTGCCATA	ACTGGAGCCC	GAGAGTTTGG	350
	L T S N	E R S	A A I	T G A	R E F G	
	CAATCTCTGG	TATCTCAGTC	AGGTCAATGA	TAGGATGACA	ACGGAGGAA	400
	N L W	Y L S Q	V N D	R M T	T E E R	
	GAGACGATT	CCCCACAGG	CAGCAGGCAG	TTCACAGTGT	AGCTCTCAT	450
	E R F	P T G	Q Q A V	P S V	A P H	
	TGGACACAG	AATCAGATCA	TGGAGATTGG	TGCCGCAGAC	ATTTACAAC	500
	W D T E	S E H	G D W	C R R H	L Q L	
	TGGTGCTAN	AAGGACTING	GAAACTTAGG	AAGACTANGA	ATTATTCAN	550
	A C X	K D X G	K L G	R L X	I I Q X	
	GATGTCCACT	ANNACACAGG	GGAAAGGAAG	AAAATCCTAC	TGCTTTCTTG	600
	C P L	X H R	G K E E	N P T	A F L	
	GAGAGCTAA	GGGAGGCATT	GAGGAAGCAT	ACCAGGCAAG	TGGACATTGG	650
	E R L R	E A L	R K H	T R Q V	D I G	
	AGGCTCTGGA	AAAGGGAAAA	GTGGGGCAAA	TTATATGCTT	AATAGGGCTT	700
	G S G	K G K S	W A N	Y M P	N R A C	
	GCTTCCAGTG	CAGTCTACAA	GGAGGCTTGA	GAAAAGATTG	TCCAAGTAGA	750
	F Q C	S L Q	G R F R	K D C	P S R	
	AATAAGGCGC	CCCTGGTCCA	TGCCCCCTTAT	GTCAGGGGAA	TCACTGGAG	800
	N K P P	L V H	A P Y	V K G I	T G R	
	GCTTACTGOC	CCAGGGGAAG	AAGGTCTCTT	GAGTCAGAAG	CCACTAACCT	850
	P T A	P G D E	G P L	S Q K	P L T	
GA						852

FIG 37

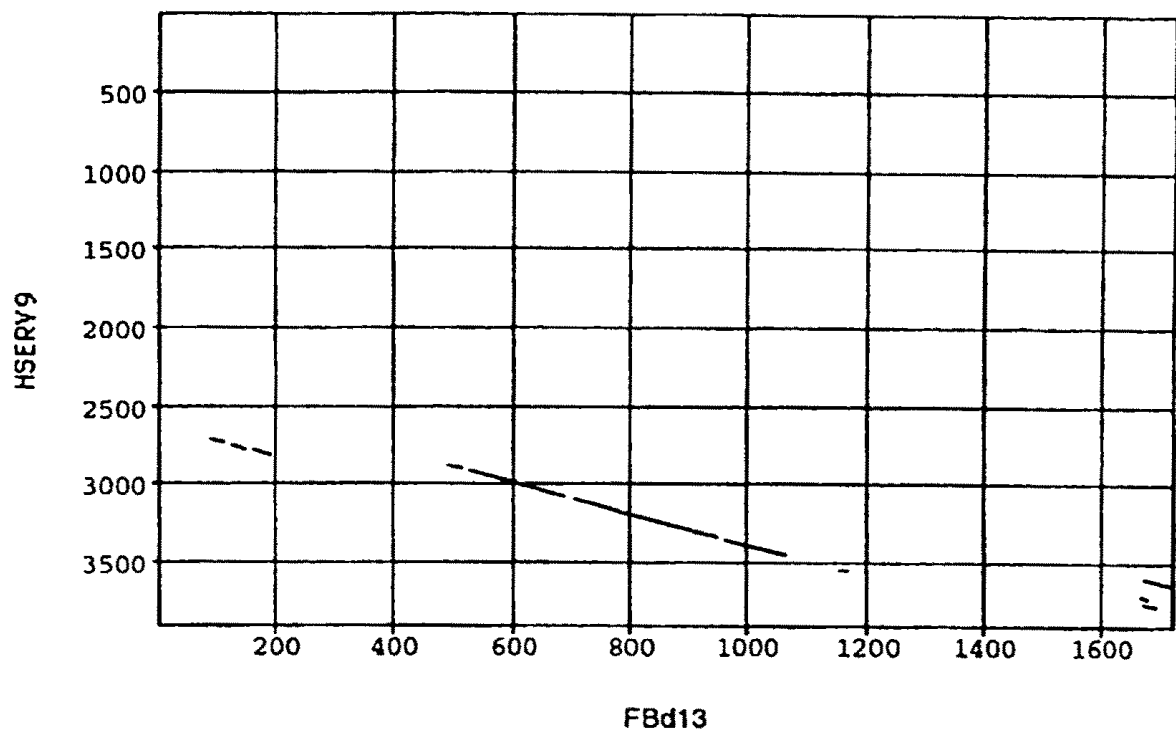


FIG 38
a

	10	20	30	40	50	
	1234567890	1234567890	1234567890	1234567890	1234567890	
	AACGAAACTC	AGAAAGCCAA	TACCCATTTA	GTAAGATGGA	CACCAGAAGC	50
	K E T Q	K A N	T H L	V R W T	P E A	
	R K L	R K P I	P I .	. D G	H Q K Q	
	G N S	E S Q	Y P F S	K M D	T R S	
	AGAAGCAGCT	TTCAGGCCC	TAAAGAAATC	CCTAAGCCAA	GGCCAGTGT	100
	E A A	F Q A L	K K S	L T Q	A P V L	
	K Q L	S R P	. R' N P	. P K	P Q C	
	R S S F	P G P	K E I	P N P S	P S V	
	TAAGCTTGGC	AACGGGCCAA	GACTTTTCTT	TATATGTCAC	AGAAAAACAG	150
	S L P	T G Q	D F S L	Y V T	E K Q	
	. A C Q	R G K	T F L	Y M S Q	K N R	
	K L A	N G A R	L F F	I C H	R K T G	
	GAATAGCTCT	AGGAGTCCTT	ACACAGGTCC	AAGGGACAAG	CTTGCAACTT	200
	E . L .	E S L	H R S	K G Q A	C N L	
	N S S	R S P Y	T G P	R D K	L A T C	
	I A L	G V L	T Q V Q	G T S	L Q P	
	GTGGCATACC	TGAGTAAGGA	AACGATGTA	NTGGCAAAGG	GTTGGGCTCA	250
	W H T	. V R K	L M X	W Q R	V G L I	
	G I P	E . G N	. C X	G K G	L A S	
	V A Y L	S K E	T D V	X A K G	W P H	
	TTGTTTACAG	GTAGGCAGC	AGTAGCAGTC	TTAGTTTCTG	AAACAGTTAA	300
	V Y R	. G S	S S S L	S F .	N S .	
	L F T G	R A A	V A V	L V S E	T V K	
	C L Q	V G Q Q	. Q S	. F L	K Q L K	
	AATAATACAG	GGAGAGATC	TTACTGTGTG	GACATCTCAT	GATGTGAACG	350
	N N T G	K R S	Y C V	D I S .	C E R	
	I I Q	G R D L	T V W	T S H	D V N G	
	. Y R	E E I	L L C G	H L M	M . T	
	GCATACTCAC	TGCTAAAGAG	GACTTGTGGC	TGTCAGACAA	CCATTACTT	400
	H T H C	. R G	L V A	V R Q	P F T	
	I L T	A K E	D L W L	S D N	H L L	
	A Y S L	L K R	T C G	C Q T T	I Y L	
	AAATAGCAGG	TTCTATTACT	TGAAGTCCA	GIGCTGGAC	TGCACATTG	450
	I A G	S I T	. S A S	A A T	A H L	
	K . Q V	L L L	E V P	V L R L	H I C	
	N S R	F Y Y L	K C Q	C C D	C T F V	
	TGCAACTCTT	AACCCAGCCA	CATTTCCTCC	AGACAATGAA	GAAAAGATAG	500
	C N S .	P S H	I S S	R Q .	R K D R	
	A T L	N P A T	F L P	D N E	E K I E	
	Q L L	T Q P	H F F Q	T M K	K R .	

FIG38
b

	10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	
AACATAACTG	TCAACAAGTA	ATTGCTCAAA	CCTATGCTGC	TCGAGGGGAC		550
T . L	S T S N	C S N	L C C	S R G P		
H N C	Q Q V	I A Q T	Y A A	R G D		
N I T V	N K .	L L K	P M L L	E G T		
CTTCTAGAGG	TTCCTTGAC	TGATCCCGAC	CTCAACTTGT	ATACTGATGG		600
S R G	S L D .	S R P	Q L V	Y . W		
L L E V	P L T	D P D	L N L Y	T D G		
F . R	F P .	L I P T	S T C	I L M E		
AAGTTCTTIG	GCAGAAAAAG	GACTTTGAAA	AGCGGGGTAT	GCAGTGATCA		650
K F L G	R K R	T L K	S G V C	S D Q		
S S L	A E K G	L . K	A G Y	A V I S		
V P W	Q K K	D F E K	R G M	Q . S		
GIGATAATGG	AATACTTGAA	AGTAATCGCC	TCACTCCAGG	AACTAGTGCT		700
. . W	N T . K	. S P	H S R	N . C S		
D N G	I L E	S N R L	T P G	T S A		
V I M E	Y L K	V I A	S L Q E	L V L		
CACCTGGCAG	AACTAATAGC	CCTCACTTGG	GCACTAGAAT	TAGGAGAAGG		750
P G R	T N S	P H L G	T R I	R R R		
H L A E	L I A	L T W	A L E L	G E G		
T W Q	N . .	P S L G	H . N	. E K E		
AAAAAGGGTA	AATATATATT	CAGACTCTAA	GTATGCTTAC	CTAGTCTTCC		800
K K G K	Y I F	R L .	V C L P	S P P		
K R V	N I Y S	D S K	Y A Y	L V L H		
K G .	I Y I	Q T L S	M L T	. S S		
ATGCCCATGC	AGCAATATGG	AGAGAGAGGG	AATTCTTAAC	TTCTGAGGGA		850
C P C	S N M E	R E G	I P N	F . G N		
A H A	A I W	R E R E	F L T	S E G		
M P M Q	Q Y G	E R G	N S .	L L R E		
ACAACCTATCA	ACCATCAGGG	AAGCCATTAG	GAGATTATTA	TTGGCTGTAC		900
T Y Q	P S G	K P L G	D Y Y	W L Y		
T P I N	H Q G	S H .	E I I I	G C T		
H L S	T I R E	A I R	R L L	L A V Q		
AGAAACCTAA	AGAGGTGGCA	GTCTTACACT	GCCAGGGTCA	TCAGGAAGAA		950
R N L K	R W Q	S Y T	A R V I	R K K		
E T .	R G G S	L T L	P G S	S G R R		
K P K	E V A	V L H C	Q G H	Q E E		
GAGGAAAGGG	AAATAGAAGG	CAATGGCCAA	GCGGATATTG	AAGCAAAAAA		1000
R K G K	. K A	I A K	R I L	K Q K K		
G K G	N R R	Q S P S	G Y .	S K K		
E E R E	I E G	N R Q	A D I E	A K K		

FIG 38
C

10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	
AGCCGCAAGG	CAGGACTCTC	CATTAGAAAT	GCTTATAGAA	GGACCCCTAG	1050
P Q G	R T L H .	K C L .	K D P .		
S R K A	G L S I R N	A Y R R	T P S		
A A R	Q D S P L E M	L I E	G P L V		
TATGGGGTAA	TCCCCTCTGG	GAAACCAAGC	CCAGTACTC	AGCAGGAAAA	1100
Y G V I	P S G K P S	P S T Q	Q E K		
M G .	S P L G N Q A	P V L	S R K N		
W G N	P L W E T K P	Q Y S	A G K		
ATAGAATAGG	AAACCTCACA	AGGACATACT	TTCTCCCCCT	CCAGATGGCT	1150
. N R	K P H K D I L	S S P	P D G .		
R I G	N L T R T Y F	P P L	Q M A		
I E .	E T S Q G H T	F L P S	R W L		
AGCCACTGAG	GAAGGAA				1167
P L R	K E				
S H .	G R				
A T E	E G				

FIG 39

	10	20	30	40	50	
	1234567890	1234567890	1234567890	1234567890	1234567890	
	AACTTGCGTG CTAGAAGGAC TAAGGAAAAC TAGGAAGACT ATGAATTATT					50
	N L R A R R T K E N . E D Y E L F					
	T C V L E G L R K T R K T M N Y S					
	L A C . K D . G K L G R L . I I					
□	CAATGATGTC CACTATAACA CAGGGGAAAG GAAGAAAATC CTACTGCCTT					100
	N D V H Y N T G E R K K I L L P F					
	M M S T I T Q G K G R K S Y C L					
	Q . C P L . H R G K E E N P T A F					
	TCTGAGAGA CTAGGGAGG CATTGAGGAA GCATACCAGG CAAGTGACA					150
	W R D . G R H . G S I P G K W T					
	S G E T K G G I E E A Y Q A S G H					
	L E R L R E A L R K H T R Q V D I					
	TTGAGGCTC TGGAAAAGG AAAAGTTGG CAAATTGAAT GCCTAATAGG					200
	L E A L E K G K V G Q I E C L I G					
	W R L W K R E K L G K L N A . . G					
	G G S G K G K S W A N . M P N R					
	GCTTCTTCC AGTGCGTCT ACAAGGAGC TTTAGAAAAG ATTGTCCAAG					250
	L A S S A V Y K D A L E K I V Q V					
	L L P V Q S T R T L . K R L S K					
	A C F Q C S L Q G R F R K D C P S					
	TAGAAATAAG CCGCCCTCG TCCATGCCCC TTATGTCAAG GGAATCACTG					300
	E I S R P S S M P L M S R E S L					
	. K . A A P R P C P L C Q G N H W					
	R N K P P L V H A P Y V K G I T G					
	GAAGGCTAC TGGCCAGGG GACGAAGGTC CTCTGAGTCA GAAGCCACTA					350
	E G L L P Q G T K V L . V R S H .					
	K A Y C P R G R R S S E S E A T N					
	R P T A P G D E G P L S Q K P L					
	ACCTGATGAT CCAGCAGCAG GACTGAGGGT GCGCGGGCCA AGTGCCAGCC					400
	P D D P A A G L R V P G A S A S P					
	L M I Q Q Q D . G C P G Q V P A					
	T . . S S S R T E G A R G K C Q P					
	CATGCCATCA CCTCAGAGC CCGGGTATG TTTGACCATT GAGAGCCAGG					450
	C H H P Q S P G Y V . P L R A R					
	H A I T L R A P G M F D H . E P G					
	M P S P S E P R V C L T I E S Q E					
	AAGTTAACTG TCTCTGGAC ACTGGGCGAG CCTTCTCAGT CTTACTTTCC					500
	K L T V S W T L A Q P S Q S Y F P					
	S . L S P G H W R S L L S L T F L					
	V N C L L D T G A A F S V L L S					

FIG 39
b

	10	20	30	40	50	
	1234567890	1234567890	1234567890	1234567890	1234567890	
	TGTCOCAGAC	AATGTGCTC	CAGATCTGTC	ACTATCCGAG	GGTCTCTAAG	550
	V P D N C P P	D L S L S E	G S . D			
	S Q T I V L	Q I C H	Y P R	G P K		
	C P R Q L S S	R S V T I R	G V L R			
	ACAGCCAGTC	ACTACATACT	TCTCTCAGCC	ACTAAGTTGT	GACTGGGGAA	600
	S Q S L H T	S L S H .	V V T G E			
	T A S H Y I L	L S A T K L .	L G N			
	Q P V T T Y F	S Q P L S C	D W G T			
	CTTTACTCTT	TTACATGCT	TTTCTAATTA	TGCTTGAAAG	CCCCACTCCC	650
	L Y S F H M L F .	L C L K A P L P				
	F T L F T C F	S N Y A .	K P H S L			
	L L F S H A	F L I M P E S	P T P			
	TGTTAGGGA	GAGACATTTT	AGCAAAAGCA	GGGGCCATTA	TACACCTGAA	700
	C . G E T F .	Q K Q G P L	Y T . T			
	V R E R H F	S K S R G H Y	T P E			
	L L G R D I L	A K A G A I I	H L N			
	CATAGGAAA	GGAATACCA	TTTGCTGTCC	OCTGCTTGAG	GAAGGAATTA	750
	. E K E Y P	F A V P C L R	K E L			
	H R K R N T H	L L S P A .	G R N .			
	I G K G I P I	C C P L L E	E G I N			
	ATGCTGAAGT	CTGGGCAATA	GAAGGACAAT	ATGGACAAGC	AAAGAATGCC	800
	I L K S G Q .	K D N M D K Q	R M P			
	S . S L G N R	R T I W T S	K E C P			
	P E V W A I	E G Q Y G Q A	K N A			
	CGTCTGTTC	AAGTTAACT	AAAGGATTCT	GCTCTCTTTC	CCTACCAAG	850
	V L F K L N .	R I L P P F	P T K G			
	S C S S .	T K G F C L L S	L P K			
	R P V Q V K L	K D S A S F P	Y Q R			
	GAAGTACCCT	CTTAGACCCG	AGGCCCTACA	AGGACTCAAA	AGATTGTGTA	900
	S T L L D P	R P Y K D S K	D C .			
	E V P S .	T R G P T R T Q K	I V K			
	K Y P L R P E	A L Q G L K	R L L R			
	GGACCTAAAA	GOCCAAGGCC	TAGTAAAACC	ATGCAGTAGC	COCTGCAATA	950
	G P K S P R P	S K T M Q .	P L Q Y			
	D L K A Q G L	V K P C S S	P C N T			
	T . K P K A .	. N H A V A	P A I			
	CTCCAATTTT	AGGAGTAAGG	AAACCCCAAG	GACAGTGGAG	GTTAGTGCAA	1000
	S N F R S K E	T Q R T V E	V S A R			
	P I L G V R	K P N G Q W R	L V Q			
	L Q F . E .	G N P T D S G G	. C K			

FIG 39
C

	10	20	30	40	50	
	1234567890	1234567890	1234567890	1234567890	1234567890	
	GATCTCAGGA	TTATTAAATGA	GGCTGTTTTT	CCTCTATAACC	CAGCTGTATC	1050
	S Q D Y . .	G C F S	S I P	S C I		
	D L R I	I N E	A V F	P L Y P	A V S	
	I S G	L L M R	L F F	L Y T	Q L Y L	
	TAGCCCTTAT	ACTCTGCTTT	CCCTAATAACC	AGAGGAAGCA	GAGTAGTTTA	1100
	. P L Y	S A F	P N T	R G S R	V V Y	
	S P Y	T L L S	L I P	E E A	E . F T	
	A L I	L C F	P . Y Q	R K Q	S S L	
	CAGTCCTGGA	CCTTAAGGAT	GGCTCTTTCT	GCATCCCTGT	ACATCCTGAT	1150
	S P G	P . G C	L F L	H P C	T S . F	
	V L D	L K D	A S F C	I P V	H P D	
	Q S W T	L R M	P L S	A S L Y	I L I	
	TCTCAATTCT	TGTTTGTCCT	TGAAGATCCT	TIGAACCCTAA	TGTCCTAATT	1200
	S I L	V C L	. R S F	E P N	V S I	
	S Q F L	F V F	E D P	L N P M	S Q F	
	L N S	C L S L	K I L	. T Q	C L N S	
	CACCTGGACT	GTTTTACCCC	AGGGGTTCGG	GGATAGCCCC	CATCTATTIG	1250
	H L D C	F T P	G V P	G . P P	S I W	
	T W T	V L P Q	G F R	D S P	H L F G	
	P G L	F Y P	R G S G	I A P	I Y L	
	GCCAGGCATT	AGCCCAAGAC	TTGAGCCAAT	TCTCATACTT	GGACATCTTG	1300
	P G I	S P R L	E P I	L I P	G H L V	
	Q A L	A Q D	L S Q F	S Y L	D I L	
	A R H .	P K T	. A N	S H T W	T S C	
	TCCTTGGTA	TGGATGATT	TAATTTTAGC	CACCCGTTC	GAAACCTTGT	1350
	L R Y	G M I	. F . P	P V Q	K P C	
	S F G M	G . F	N F S	H P F R	N L V	
	P S V	W D D L	I L A	T R S	E T L C	
	GCCATCAAGC	CACCCAAGCG	TTCCTAAATT	TCCTCACTCC	GTGTGGCTAC	1400
	A I K ' P	P K R S	. I S S L R	V A T		
	P S S	H P S V	L K F	P H S	V W L Q	
	H Q Å	T Q A	F L N F	L T P	C G Y	
	AAGGTTTCCA	AACCAAAGGC	TCAGCTCTGC	TCACAGCAGG	TTAAATACTT	1450
	R F P	N Q R L	S S A	H S R	L N T .	
	G F Q	T K G	S A L L	T A G	. I L	
	K V S K	P K A	Q L C	S Q Q V	K Y L	
	AGGTTTAAAA	TTATCCAAG	GCACCAAGGC	CCTCTGTGAG	GAATGTATCC	1500
	G . N	Y P K	A P G P	S V R	N V S	
	R V K I	I Q R	H Q G	P L .	G M Y P	
	G L K	L S K G	T R A	L C E	E C I Q	

FIG 39
d

10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	
AACCTGTACT GGCTTATCTT CATCCCAAAA CCTAAAGCA ACTAAGAAGG					1550
N L Y W	L I F I P K	P . S N	. E G		
T C T	G L S S S Q N	P K A	T K K V		
P V L	A Y L H P K	T L K Q	L R R		
TCCTTGGCAT AACAGGTTTC TGCCGAA					1577
P W H	N R F L	P			
L G I	T G F C R				
S L A .	Q V S A E				

FIG 40

10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	
TCCAGCAGCA	GGACTGAGGG	TGCCCCGGGGC	AAGTGCCAGC	CCATGCCATC	50
S S S R	T E G	A R G	K C Q P	M P S	
ACCCTCAGAG	CCCCGGGTAT	GTTTGACCAT	TGAGAGCCAG	GAAGTTAACT	100
P S E	P R V C	L T I	E S Q	E V N C	
GTCTCTGGA	CACTGGCGCA	GCCTTCTCAG	TCTTACTTTC	CTGTCCCAGA	150
L L D	T G A	A F S V	L L S	C P R	
CAATTGTCTT	CCAGATCTGT	CACTATCCGA	GGGGTCTTAA	GACAGCCAGT	200
Q L S S	R S V	T I R	G V L R	Q P V	
CACTACATAC	TTCTCTCAGC	CACTAAGTTG	TGACTGGGGA	ACTTTACTCT	250
T T Y	F S Q P	L S C	D W G	T L L F	
TTTCACATGC	TTTTCTAATT	ATGCCTGAAA	GCCCCACTCC	CTTGTTAGGG	300
S H A	F L I	M P E S	P T P	L L G	
AGAGACATTT	TAGCAAAAGC	AGGGGCCATT	ATACACCTGA	ACATAGGAAA	350
R D I L	A K A	G A I	I H L N	I G K	
AGGAATACCC	ATTTGCTGTC	CCCTGCTTGA	GGAAGGAATT	AATCTGAAG	400
G I P	I C C P	L L E	E G I	N P E V	
TCTGGGCAAT	AGAAGGACAA	TATGGACAAG	CAAAGAATGC	CCGTCTGTGT	450
W A I	E G Q	Y G Q A	K N A	R P V	
CAAGTTAAAC	TAAAGGATTC	TGCCCTCCTTT	CCCTACCAAA	GGAAGTACCC	500
Q V K L	K D S	A S F	P Y Q R	K Y P	
TCTTAGAACC	GAGGCCCTAC	AAGGACTCAA	AAGATTGTTA	AGGACCT	547
L R P	E A L Q	G L K	R L L	R T	

FIG 41

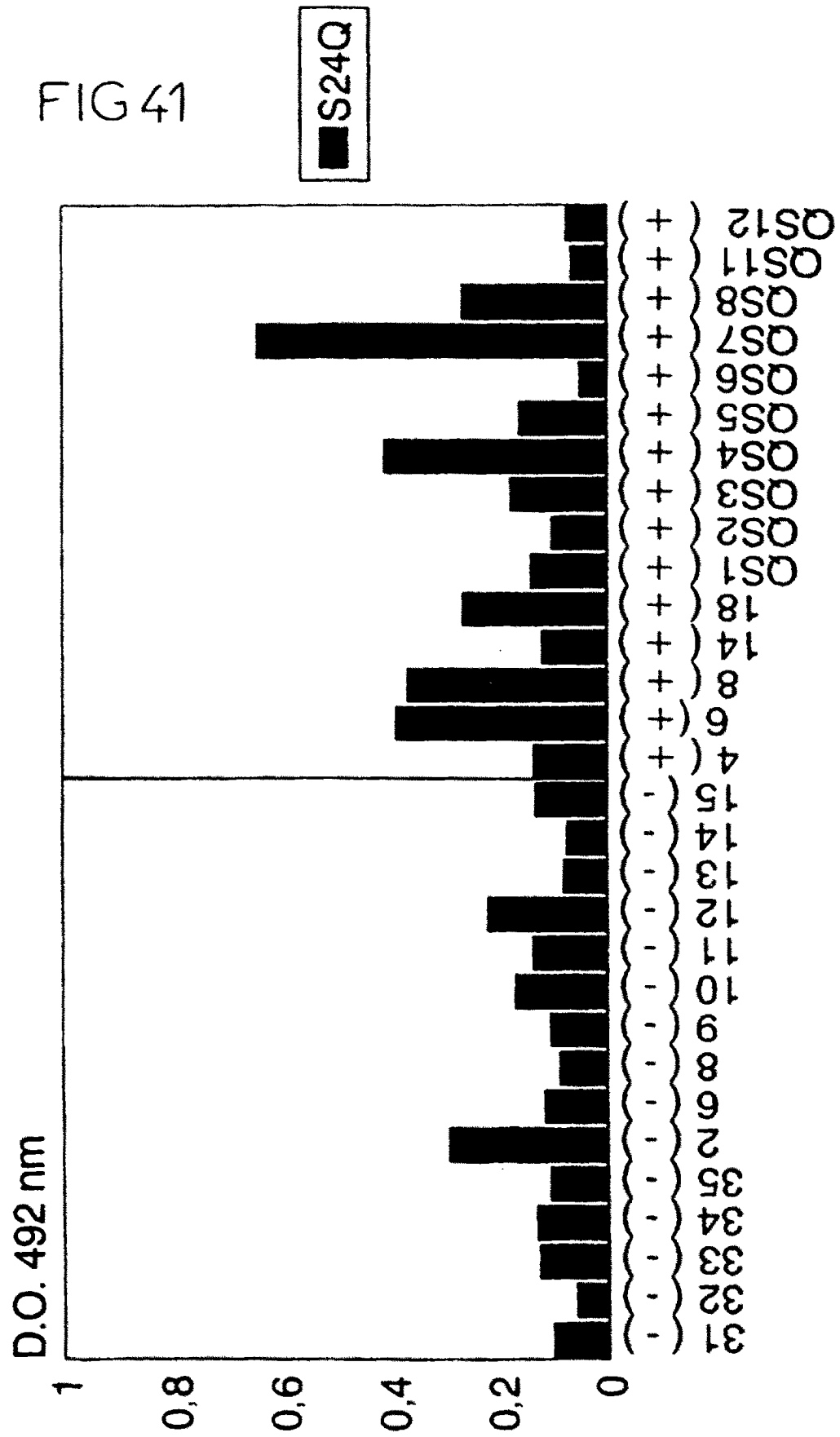


FIG 42

