

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 327 217**

51 Int. Cl.:

C12N 7/00

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA MODIFICADA
TRAS OPOSICIÓN

T5

96 Fecha de presentación y número de la solicitud europea: **04.04.2002** **E 02759815 (0)**

97 Fecha y número de publicación de la solicitud europea: **28.01.2004** **EP 1383796**

54 Título: **Cepa neurovirulenta del virus West-Nile y sus aplicaciones**

30 Prioridad:

04.04.2001 FR 0104599

06.09.2001 FR 0111525

45 Fecha de publicación y mención en BOPI de la
traducción de la patente modificada:
11.02.2013

73 Titular/es:

INSTITUT PASTEUR (50.0%)

28, RUE DU DOCTEUR ROUX

75724 PARIS CEDEX 15, FR y

KIMRON VETERINARY INSTITUTE (50.0%)

72 Inventor/es:

DESPRES, PHILIPPE;

DEUBEL, VINCENT;

GUENET, JEAN-LOUIS;

DROUET, MARIE-THÉRÈSE;

MALKINSON, MERTYN;

BANET, CAROLINE;

FRENKIEL, MARIE-PASCALE;

COURAGEOT, MARIE-PIERRE;

COULIBALY, FASSÉLI;

CATTEAU, ADELINE;

FLAMAND, MARIE;

WEBER, PATRICK y

CECCALDI, PIERRE-EMMANUEL

74 Agente/Representante:

PONS ARIÑO, Ángel

DESCRIPCIÓN

Cepa neurovirulenta del virus West-Nile y sus aplicaciones

- 5 La presente invención se refiere a una cepa neuroinvasiva y neurovirulenta del virus West-Nile (*o virus del Nilo Occidental*), denominada IS-98-ST1, a moléculas de ácido nucleico procedentes de su genoma, a las proteínas y péptidos codificados por dichas moléculas de ácido nucleico así como a sus aplicaciones.

La presente invención también se refiere a todas las variantes de la cepa viral IS-98-ST1 que tienen al menos una
10 mutación en la secuencia nucleica correspondiente a NS5.

La familia de los *Flaviviridae* agrupa a los virus del género flavivirus responsables de patologías humanas graves tales como el dengue, la fiebre amarilla, encefalitis transmitidas por las garrapatas, la encefalitis japonesa, la encefalitis del Nilo Occidental y los virus de las hepatitis C y G. Si los flavivirus pueden provocar una morbilidad y
15 una mortalidad importantes en el ser humano, la infección es generalmente asintomática y solamente una fracción de los individuos infectados desarrolla una enfermedad grave.

Los flavivirus son pequeños virus con envoltura. Su genoma es una molécula de ARN monocatenario de polaridad positiva de aproximadamente 11 000 bases. El ARN genómico se asocia a varias copias de la proteína de cápside C para formar la nucleocápside; esta está rodeada por una envoltura viral constituida por una doble capa lipídica
20 procedente de las membranas del retículo endoplásmico (RE) en las que están ancladas la proteína de envoltura E y la proteína de membrana M. El ARN genómico de los flavivirus contiene un único marco de lectura abierto de aproximadamente 10.500 nucleótidos flanqueado por dos regiones cortas no codificantes en sus extremos 5' y 3'. El genoma se traduce a una poliproteína de aproximadamente 3.400 aminoácidos que es el precursor de las proteínas
25 estructurales C, prM (el precursor intracelular de M) y E en su parte N-terminal y de al menos siete proteínas no estructurales (NS) de NS1 a NS5 en su parte C-terminal.

Hasta hace muy poco, se reconocía al virus West-Nile como un virus poco patógeno, responsable de un síndrome gripal y presente en África, en el sur de Europa y en Oriente Medio; se ha aislado en el transcurso de epidemias,
30 ocurridas particularmente en Israel en los años 1950 y en Sudáfrica en los años 1970.

Muy recientemente, la epidemiología del virus West-Nile se ha modificado y se ha observado un número creciente de casos de encefalitis en el transcurso de las epidemias ocurridas en Rumanía en 1996, en Israel en 1998 y en los EE.UU. en 1999.
35

Se han aislado cepas patógenas durante estas epidemias (Anderson y col., y Lanciotti y col., Science, 1999, 286: 2331-2333, 2333-2337), en particular la cepa NY1999 (GenBank nº AF202541, Lanciotti y col., mencionado anteriormente), cuya patogenicidad se correlacionará con la presencia de un sitio de glucosilación NTS en la proteína de envoltura E (Jordan y col., Viral Immunol., 2000, 13, 4: 435-446).
40

Los factores virales mal identificados podrían ser responsables de la gravedad de la infección, mientras que la constitución genética del hospedador (humano o no humano) contribuiría a la resistencia a la infección.

Sin embargo, los datos relativos a estas cepas patógenas aisladas recientemente no han permitido determinar todos los factores virales y los genes del hospedador implicados en la sensibilidad/resistencia a la infección por los
45 *Flaviviridae*.

Los modelos murinos han permitido establecer la existencia de una resistencia genética a la infección por los flavivirus. Se ha demostrado que algunas líneas de ratón recientemente derivadas del estado silvestre y que pertenecen a las especies *Mus musculus musculus* o *Mus spretus* (Det, BSVR, BRVR, PRI, CASA/Rk y CAST/Ei)
50 son resistentes a la infección por los flavivirus, mientras que las líneas consanguíneas de laboratorio más habituales que derivan mayoritariamente de la especie *Mus musculus domesticus*, no la resisten (Sangster y col., J. Virol., 1993, 67: 340-347).

La resistencia está controlada por al menos un locus autosómico denominado *Flv*, localizado en el cromosoma 5, en el ratón y tres alelos *Flv^s*, *Flv^r* y *Flv^{mr}* otorgan respectivamente la sensibilidad, la resistencia y la resistencia intermedia a la infección por los flavivirus. Usando una cepa del flavivirus de la encefalitis del Valle de Murray y de ratones procedentes del retrocruzamiento de la línea de ratón resistente C3H/RV con las líneas de ratón sensibles C3/He o BALB/c, el locus *Flv* se ha localizado en una región de 0,9 cM del cromosoma 5, en el ratón, entre los
55

marcadores *D5Mit68* y *D5Mit242* (G.R. Shellam y col., Rev. Sci. Tech. Off. Epiz. 1998, 17: 231-248).

Los autores de la invención han aislado ahora una nueva cepa del virus West-Nile, a partir de muestras extraídas de cigüeñas en Israel (en la ciudad de Eilat) en septiembre de 1998, que se ha seleccionado para el estudio de la resistencia/sensibilidad de un hospedador (mamífero humano o no humano) a la infección por los virus de la familia de los *Flaviviridae*.

De acuerdo con la invención, dicha cepa neurovirulenta y neuroinvasiva del virus West-Nile aislada, denominada IS-98-ST1, se caracteriza porque su genoma está constituido por la secuencia SEQ ID N° 1 que codifica una poliproteína que presenta la secuencia SEQ ID N° 2.

Los autores de la invención demostraron particularmente que los ratones de laboratorio son extremadamente sensibles a la infección por la cepa IS-98-ST1 mientras que los ratones de las líneas SEG, WMP, STF y MAI que derivan de ratones silvestres que pertenecen a especies diferentes aunque del mismo género *Mus*, son completamente resistentes a la infección por esta cepa; una inoculación por vía intraperitoneal de 1.000 UFF (Unidades Formadoras de Focos; UFF: DL50 = 100) es mortal al 100% para los ratones de laboratorio, mientras que los ratones silvestres no presentan ningún síntoma; además, el virus se replica en estos ratones, como muestra la aparición de anticuerpos séricos específicos.

La presente invención también tiene por objeto reactivos, derivados de la cepa IS-98-ST1, usados para el estudio y el diagnóstico de las infecciones por *Flaviviridae*, reactivos que se seleccionan entre el grupo constituido por los siguientes reactivos:

(a) una molécula de ácido nucleico elegida entre la secuencia SEQ ID N° 1, los fragmentos de al menos 15 nucleótidos de la secuencia SEQ ID N° 1 y las secuencias complementarias de sentido y antisentido de las secuencias precedentes, con exclusión del fragmento que presenta la secuencia GENBANK AF205882.

(b) un vector recombinante que comprende una molécula de ácido nucleico tal como se define en (a),

(c) una célula transformada por una molécula de ácido nucleico tal como se define en (a), un vector tal como se define en (b) o una cepa neurovirulenta del virus West-Nile tal como se define en (a),

(b) o una cepa neurovirulenta del virus West-Nile tal como se define en (a),

(d) una proteína o un péptido codificado por una molécula de ácido nucleico tal como se define en (a),

(e) un anticuerpo policlonal, que puede obtenerse por inmunización de un mamífero no humano con la cepa IS-98-ST1 del virus West-Nile tal como se ha definido anteriormente; preferentemente, dicho mamífero no humano es un ratón homocigótico para el alelo *Flv*, resistente a la infección por *Flaviviridae*, y

(f) un anticuerpo policlonal o monoclonal, que puede obtenerse por inmunización de un mamífero no humano con un vector recombinante tal como se define en (b) o bien una proteína o un péptido, tal como se definen en (d).

Estos diferentes reactivos se preparan y usan de acuerdo con técnicas clásicas de biología molecular y de inmunología, siguiendo los protocolos convencionales tales como los descritos en Current Protocols in Molecular Biology (Frederick M. AUSUBEL, 2000, Wiley and Son Inc, Library of Congress, EEUU) y en Current Protocols in Immunology (John E. Coligan, 2000, Wiley and Son Inc. Library of Congress. EEUU).

Los fragmentos de ácidos nucleicos tales como los definidos anteriormente, en particular los correspondientes a las secuencias SEQ ID N° 3-11 se usan por ejemplo como sonda o como cebador para el diagnóstico de una infección por el virus West-Nile; la infección se detecta por ejemplo por PCR y/o por hibridación, a partir de los ácidos nucleicos extraídos de una muestra biológica tomada de un individuo que puede estar infectado o un animal de laboratorio inoculado con dicho virus.

De acuerdo con una realización ventajosa de dichos fragmentos, estos comprenden al menos 15 nucleótidos de la SEQ ID N° 1 situados entre 10 y 100 nucleótidos en la dirección 5' y en la dirección 3' de uno de los codones en la siguiente posición:

- codón Asparagina (N) en posiciones 2518-2520, correspondiente al resto en posición 17 de la proteína NS1 o en

posición 808 de la secuencia SEQ ID N° 2 (la secuencia de la proteína NS1 se extiende desde los codones en posición 792 a 1144 de la secuencia SEQ ID N° 2, correspondientes a los nucleótidos 2470 a 3528 de la secuencia SEQ ID N° 1),

- 5 - codón Arginina (R) en posiciones 4018-4020, correspondiente al resto en posición 164 de la proteína NS2A o en posición 1308 de la secuencia SEQ ID N° 2 (la secuencia de la proteína NS2A se extiende desde los codones en posición 1145 a 1374 de la secuencia SEQ ID N° 2, correspondientes a los nucleótidos 3529 a 4218 de la secuencia SEQ ID N° 1),
- 10 - codones Glicina (G) en posiciones 4462-4464 y ácido glutámico (E) en posiciones 4465-4467, correspondientes respectivamente al resto en posición 82 y 83 de la proteína NS2B o en posición 1456 y 1457 de la secuencia SEQ ID N° 2 (la secuencia de la proteína NS2B se extiende desde los codones en posición 1375 a 1505 de la secuencia SEQ ID N° 2, correspondientes a los nucleótidos 4219 a 4611 de la secuencia SEQ ID N° 1),
- 15 - codones Prolina (P) en posiciones 6097-6099 y ácido glutámico en posiciones 6172-6174, correspondientes respectivamente al resto en posición 496 y 521 de la proteína NS3 o en posición 2001 y 2026 de la secuencia SEQ ID N° 2 (la secuencia de la proteína NS3 se extiende desde los codones en posición 1506 a 2124 de la secuencia SEQ ID N° 2, correspondientes a los nucleótidos 4612 a 6468 de la secuencia SEQ ID N° 1), y
- 20 - codones Serina (S) en posiciones 7840-7842, Asparagina (N) en posiciones 8518-8520 y Alanina (A) en posiciones 8794-8796, correspondientes respectivamente al resto en posición 54, 280 y 372 de NS5 o en posición 2582, 2808 y 2900 de la secuencia SEQ ID N° 2 (la secuencia de la proteína NS5 se extiende desde los codones en posición 2529 a 3430 de la secuencia SEQ ID N° 2, correspondientes a los nucleótidos 7681 a 10386 de la secuencia SEQ ID N° 1).
- 25 Dichos cebadores son útiles para amplificar fragmentos que contienen dichos codones.

Dichos cebadores están situados entre 10 y 100 nucleótidos en la dirección 5' o en la dirección 3' de dichos codones.
- 30 De acuerdo con otra realización ventajosa de dichos fragmentos, estos están constituidos por los fragmentos que comprenden los codones mencionados anteriormente, preferentemente entre 50 y 200 nucleótidos, que se amplifican usando los cebadores tales como se han definido anteriormente.
- 35 Los autores de la invención también describen fragmentos que comprenden al menos 15 nucleótidos de la SEQ ID N° 1 en la dirección 5' o en la dirección 3' del codón Alanina (A) en posiciones 1117-1119, correspondientes al resto en la posición 51 de la proteína E o en la posición 341 de la secuencia de la poliproteína viral de la secuencia SEQ ID N° 2 (la secuencia de la proteína E se extiende desde los codones en posición 291 a 791 de la secuencia SEQ ID N° 2, correspondientes a los nucleótidos 967 a 2469 de la secuencia SEQ ID N° 1).
- 40 Los vectores recombinantes tales como los definidos anteriormente, en particular los vectores de expresión y las células transformadas por dichos vectores de expresión se usan ventajosamente para la producción de las proteínas y de los péptidos correspondientes.
- 45 Dichas proteínas y dichos péptidos, que son adecuados para ser reconocidos y/o para inducir la producción de anticuerpos específicos del virus West-Nile, en particular de cepas neurovirulentas, son útiles para el diagnóstico de una infección por un virus West-Nile; la infección se detecta mediante una técnica apropiada, particularmente EIA, ELISA, RIA, inmunofluorescencia a partir de una muestra biológica extraída de un individuo que puede estar infectado o un animal de laboratorio inoculado con dicho virus. Las proteínas y los péptidos tales como los definidos
- 50 anteriormente también se usan para la búsqueda de socios celulares de estas proteínas virales que pueden estar implicadas en la patogenicidad (neurovirulencia) del virus West-Nile; estos socios se identifican mediante técnicas de inmunoafinidad, por ejemplo por cromatografía en columna de inmunoafinidad.

Los anticuerpos de acuerdo con la invención son útiles para el diagnóstico de una infección por un virus West-Nile,
- 55 en particular de las cepas neurovirulentas; la infección se detecta mediante una técnica apropiada, particularmente EIA, ELISA, RIA, inmunofluorescencia a partir de una muestra biológica extraída de un individuo que puede estar infectado o un animal de laboratorio inoculado con dicho virus. Entre estos, los anticuerpos producidos mediante inmunización de ratones FIV/FIV' con la cepa IS-98-ST1 poseen ventajosamente un valor cuantitativo elevado y una especificidad muy grande por el virus West-Nile.

Las células transformadas de acuerdo con la invención, en particular las células neurales (neuronas y células endoteliales) infectadas por una cepa neurovirulenta tal como se ha definido anteriormente, se usan para identificar los genes procedentes de estas células cuya expresión podría modularse en el transcurso de la infección vírica; estos genes se detectan por ejemplo mediante la tecnología de biochips, de acuerdo con los protocolos clásicos tales como los descritos en Atlas Mouse Arrays (#membranes) ATLAS™ NYLON cDNA EXPRESSION ARRAYS (CLONTECH, EEUU).

La presente invención también se refiere a un modelo de estudio no humano de la sensibilidad/resistencia a la infección por un virus de la familia de los *Flaviviridae*, **caracterizado porque** comprende al menos una cepa neurovirulenta del virus West-Nile tal como se ha definido anteriormente.

De acuerdo con una realización ventajosa de dicho modelo, este comprende además un ratón homocigótico para el alelo *Flv*^r o *Flv*^s.

La presente invención también tiene por objeto un procedimiento de detección de una infección por un virus West-Nile, **caracterizado porque** comprende:

- la amplificación de los ARN procedentes de una muestra biológica que se someterá a ensayo con ayuda de los cebadores tales como se han definido anteriormente, y

- la secuenciación del producto de amplificación obtenido.

Dicha detección puede permitir ventajosamente el pronóstico de la gravedad de una encefalitis vírica por virus West-Nile.

La cepa neurovirulenta del virus West-Nile de acuerdo con la invención se usa para el cribado de genes celulares implicados en la resistencia de un mamífero a la infección por un virus de la familia de los *Flaviviridae*, preferentemente el virus de la hepatitis C.

De manera ventajosa, dicho procedimiento de cribado comprende las siguientes etapas:

- cultivo de células derivadas de un hospedador (humano o no humano) seleccionado por su resistencia o su sensibilidad a la infección por un *Flaviviridae*,
- infección *in vitro* de dichas células por un *Flaviviridae*, y
- detección de genes expresados de manera diferencial en dichas células infectadas.

De acuerdo con la invención, dicha detección puede comprender el establecimiento del perfil de transcritos o de proteínas a partir de dichas células.

La presente invención también tiene por objeto el uso del modelo tal como se ha definido anteriormente para la clasificación de moléculas activas contra una infección vírica debida a un virus de la familia de los *Flaviviridae*.

La presente invención se refiere además a un procedimiento de clasificación de moléculas activas contra una infección por un Flavivirus, caracterizado por:

- la puesta en contacto de un cultivo de células eucariotas, procedentes de un mamífero (humano o no humano) sensible a la infección por un *Flaviviridae* con una suspensión viral de la cepa de acuerdo con la reivindicación 1, en presencia o en ausencia de la molécula que se va a someter a ensayo, y

- la detección de la amplificación/replicación del virus, mediante cualquier procedimiento conocido (cuantificación del genoma, ARNm, -proteínas, partículas virales).

La presente invención tiene también por objeto una variante de la cepa viral tal como se ha definido anteriormente, **caracterizada porque** su genoma comprende al menos una mutación en la secuencia nucleotídica correspondiente a la proteína NS5.

Además de las disposiciones anteriores, la invención también comprende otras disposiciones, que se verán en la siguiente descripción, que se refiere a ejemplos de realización del objeto de la presente invención, con referencia a

los dibujos adjuntos en los que:

- las figuras 1A a 1E representan la comparación de la secuencia de aminoácidos de las proteínas virales de la cepa IS-98-ST1 (SEQ ID N° 2), aislada a partir de cigüeñas (CI) y de la cepa New York (NY99; Genbank AF196835) aislada a partir de flamencos rosa (FLA), durante la epidemia de 1999 en los Estados Unidos,
- la figura 2 representa la cinética de mortalidad y la cinética de aparición de los anticuerpos séricos específicos en ratones sensibles *Flv^s/Flv^s* (BALB/c), infectados por la cepa IS-98-ST1 del virus West-Nile,
- 10 - la figura 3 representa la cinética de propagación de la cepa IS-98-ST1 en el sistema nervioso central de los ratones sensibles *Flv^s/Flv^s* (BALB/c),
- las figuras 4 (A, B y C) representan la cinética de aparición de los antígenos virales en las células Neuro 2a y las neuronas primarias de ratones sensibles (BALB/c) infectados por el virus West-Nile (cepa IS-98-ST1),
- 15 - la figura 5 representa la muerte por necrosis de las células Neuro 2a infectadas por el virus West-Nile (cepa IS-98-ST1),
- la figura 6 representa el protocolo experimental usado para precisar la localización del locus *Flv* en el cromosoma 5 del ratón,
- 20 - la figura 7 representa el mapa genético del locus *Flv*, determinado a partir de ratones sensibles, procedentes del primer retrocruzamiento entre las líneas resistentes (MAI/Pas y MBT/Pas) y las líneas sensibles (C57BL/6 o BALB/c). Los cuadros blancos representan los alelos BALB/c o C57BL/6 y los cuadros negros representan los alelos MAI/Pas o MBT/Pas,
- 25 - la figura 8 representa el mapa genético del locus *Flv* determinado a partir de los ratones resistentes y sensibles, procedentes del primer retrocruzamiento (BC1) entre las líneas resistentes (MAI/Pas y MBT/Pas) y las líneas sensibles (C57BL/6 y BALB/c). Las líneas grises representan los alelos (BALB/c o C57BL/6) y las líneas negras representan los alelos MAI/Pas o MBT/Pas,
- 30 - la figura 9 representa el mapa genético y el mapa físico del locus *Flv* y la posición del gen OAS en este locus, y
- la figura 10 representa la distribución de los alelos *Flv* en ratones resistentes y sensibles procedentes del primer retrocruzamiento (BC1) entre las líneas resistentes (MAI/Pas y MBT/Pas) y las líneas sensibles (C57BL/6 y BALB/c).
- 35

Ejemplo 1: Aislamiento, amplificación, purificación y valoración cuantitativa de la cepa neuroinvasiva del virus West-Nile IS-98-ST1

- 40 Se obtuvo un aislado del virus West-Nile (WN) a partir del sistema nervioso central de una cigüeña que manifestaba trastornos neuropatológicos graves, en septiembre de 1998, en Eilat (Israel). La infección de células VERO por este aislado es citolítica y la inmunofluorescencia indirecta con un líquido ascítico de ratón inmuno específico del virus West-Nile (suero inmune de referencia WN 8907) es positiva al 100%. El virus producido en células VERO se recogió y amplificó en células de mosquitos AP61 (Després y col., Virol., 1993, 196: 209-219).
- 45 El pase 1 (o P1) del virus WN por células AP61 se recogió 3 días después de la infección; posee un valor cuantitativo de $2,5 \times 10^8$ UFF/ml (Unidad Formadora de Foco) mediante la técnica de valoración cuantitativa en células AP61 descrita en Després y col. (mencionado anteriormente). El inóculo P1 del virus WN en células AP61 se identificó como la cepa IS-98-ST1.
- 50 Se obtuvo un P2 a partir de células AP61 infectadas por la cepa IS-98-ST1. P1 (valor cuantitativo: 6×10^7 UFF/ml). El inóculo P2 de IS-98-ST1 se usa para las pruebas de sensibilidad a la infección vírica en ratones adultos.
- Una inóculo viral P3 de la cepa IS-98-ST1 con un valor cuantitativo de 5×10^7 UFF/ml se produjo en células AP61.
- 55 Una preparación viral altamente purificada, preparada de acuerdo con el protocolo de purificación de flaviviriones descrito en Després y col., 1993) se obtuvo a partir de 20 placas de 150 cm^2 de células AP61 recogidas 3 días después de la infección por el inóculo P3 del virus WN cepa IS-98-ST1 (multiplicidad de infección de 0,4). La cepa IS-98-ST1 purificada en gradientes de sacarosa tiene un valor cuantitativo final de 2×10^{10} UFF/ml. Los ARN extraídos de este virus purificado se usan para amplificar los ADNc correspondientes a las proteínas virales C, prM y

NS1 o a las secuencias no codificantes en los extremos 5' y 3' del genoma viral.

Ejemplo 2: Secuenciación del genoma de la cepa neuroinvasiva IS-98-ST1

5 El genoma viral se extrajo a partir del sobrenadante del cultivo de las células VERO infectadas del ejemplo 1 con ayuda del kit "QIAamp Viral RNA" (QIAGEN), siguiendo las instrucciones del fabricante. Se amplificaron 6 productos de RT-PCR solapantes a partir de estos ARN usando los cebadores descritos por Lanciotti y col. (mencionado anteriormente). Los extremos 5' y 3' del genoma viral se amplificaron respectivamente con ayuda de los siguientes cebadores:

10

- 5' AGTAGTTCGCCTGTGTGAGCTGACAAAC 3' (SEQ ID N° 5), y
- 5' AGATCCTGTGTTCTCGCACCACCAGCCAC 3' (SEQ ID N° 6).

El ADNc que corresponde a la proteína viral E se amplificó con ayuda de los cebadores 5' GGATGGATGCT(A/T)GG(G/T)AGCAAC 3' (SEQ ID N° 7) y 5' CCATCCAAGCCTCCACATC 3' (SEQ ID N° 8), que hibridan respectivamente en el gen de la proteína M (posiciones 889 a 909 de la secuencia SEQ ID N° 1) y en el gen de la proteína NS1 (posiciones 2539 a 2557 de la secuencia SEQ ID N° 1).

Los ADNc obtenidos se purificaron por cromatografía de intercambio iónico y se precipitaron en 2 volúmenes de isopropanol. A continuación se secuenciaron los ADNc en las dos cadenas usando el kit "Taq Dye Deoxy Terminator Cycle Sequencing" (PERKIN ELMER CORP./APPLIED BIOSYSTEM) y los cebadores se separaron por 400 pares de bases en el genoma viral (Lanciotti y col., mencionado anteriormente). La secuenciación se realizó con 0,2 pmoles de ADNc purificado y 30 pmoles de cebadores, siguiendo el protocolo recomendado por el fabricante. El alineamiento de las secuencias se realiza con ayuda del programa CLUSTAL W.

25

La secuencia genómica completa de la cepa IS-98-ST1 del virus West-Nile corresponde a la secuencia SEQ ID N° 1.

El alineamiento de las secuencias de aminoácidos de la cepa IS-98-ST1 (SEQ ID N° 2) y de la cepa NY99, presentada en la figura 1, muestra que la cepa IS-98-ST1 aislada en Israel en 1998 y la cepa NY-99 aislada en Nueva York en 1999 son muy próximas (divergencia de menos del 0,2% en las secuencias de aminoácidos).

30

Sin embargo, las diferencias observadas en la cepa IS-98-ST1, respectivamente en las proteínas E (A₅₁), NS1 (N₁₇), NS2A (R₁₆₄), NS2B (G₈₂, E₈₃), NS3 (P₄₉₆, E₅₂₁) y NS5 (S₅₄, N₂₈₀, A₃₇₂) son potencialmente responsables de la neurovirulencia y de las propiedades neuroinvasivas observadas con esta cepa y pueden servir como marcador de virulencia del virus West-Nile.

35

Ejemplo 3: Clonación de las proteínas de la cepa neuroinvasiva IS-98-ST1 y usos de los plásmidos recombinantes obtenidos.

40 1- La proteína C

El ARN genómico extraído de los viriones IS-98-ST1 purificados en gradientes de sacarosa descritos en el ejemplo 1, con ayuda de la solución RNA PLUS 2 (Q. BIOGEN), se usa como matriz para amplificar la secuencia que codifica la proteína C (aminoácidos 1 a 123) mediante la técnica RT-PCR (kit Titan One Tube RT-PCR; Roche Biochemicals #1939 823).

45

El par de cebadores usado en la matriz de ARN es el siguiente:

- 5'C/WNV (secuencia de nt 81-117 de la secuencia SEQ ID N° 1)
5' TAG CAC GAA GAA TTC GAT GTC TAA AAA CCA GGA GGG 3' (SEQ ID N° 11) que contiene el sitio de restricción *EcoRI*, y
- 3'C/WNV (secuencia anti-sentido de nt 433 a 482 de la secuencia SEQ ID N° 1)
5' AAGTTAGCCCGGGTTAATGCTCCTACGCTGGCGATCAGGCCAATCAGGAC 3' (SEQ ID N° 4) que contiene el sitio de restricción *SmaI*.

55

El ADNc de la proteína C de la cepa IS-98-ST1 (aminoácidos 1 a 123) del virus WN se clonó, por un lado entre los sitios *EcoRI* y *SmaI* del plásmido pCI-neo (Promega # E1841) y, por otro lado, entre los sitios *KspI* y *SmaI* del plásmido pIVEX 2.4a (Roche).

El plásmido recombinado pCI-C/WN (depositado el 21 de junio 2001 en la *Collection Nationale de Culture de Microorganismes* del Instituto Pasteur de París, 28, rue du Docteur Roux, 75724, PARIS Cedex 15 con el número 1-2688) contiene la secuencia completa del gen de la proteína C de la cepa IS-98-ST1 del virus WN entre los promotores T7 y T3. La transcripción *in vitro* de pCI-C/WN linealizada con *NheI* bajo la dependencia del promotor T3 sintetiza un ARN de aproximadamente 350 bases complementarias de la secuencia viral genómica. La ribosonda marcada con DIG (digoxigenina) se usa para la detección de los ARN virales de sentido positivo presentes en las células infectadas por el virus WN, mediante la técnica de hibridación *in situ*, de acuerdo con el protocolo descrito en Després y col. (J. Virol., 1998, 72: 823-829).

- 10 El plásmido recombinante pIVEX-C/WN se usa para la producción masiva de la proteína C (aminoácidos 1 a 123) del virus WN en lisado bacteriano (sistema RTS 500 de Roche). La proteína recombinante C producida *in vitro* posee en su extremo N-terminal una secuencia [His]₆ y el sitio de escisión reconocido por la proteasa Xa para permitir, por un lado, su purificación en columna de Ni y, por otro lado, la eliminación de los restos de histidina. La proteína C de la cepa IS-98-ST1 del virus WN producida de este modo se usa para estudios estructurales, para la búsqueda de socios celulares de esta proteína en columna de inmunoafinidad, y para la producción de anticuerpos monoespecíficos en conejo.

2- La proteína M

- 20 Los ADNc de la cepa IS-98-ST1 del virus WN que codifican la proteína M (aminoácidos 215 a 290 de la poliproteína viral) o su ectodominio de 41 aminoácidos (aminoácidos 215 a 255; acrónimo *ectoM*) se clonan:

- (1) en fase con el extremo C-terminal de la EGFP en el plásmido p[95-114]EGFP, derivado del plásmido pEGFP-N1 (Clontech) que comprende los restos 95-114 de la proteína C del virus del dengue de tipo 1 (cepa BR/90) fusionados en fase con la secuencia N-terminal de la proteína EGFP[215-290]WNV, para dar el plásmido p [95-114]EGFP[215-290]WNV,

- (2) en el plásmido pIVEX (sistema RTS 500 de Roche) para dar el plásmido pIVEX[EGFP][215-255]WNV,

- 30 (3) en el vector retroviral TRIPdeltaU3CMV, para dar el plásmido TRIPdeltaU3CMV[95-114]EGFP[215-255] WNV.

El plásmido pIVEX[EGFP][215-255]WNV permite la síntesis *acelular* y la purificación de la proteína quimérica EGFP-*ectoM* WNV que se usa, por un lado, para la producción de anticuerpos monoespecíficos dirigidos contra la proteína M del virus WN y, por otro lado, para la búsqueda de socios celulares de la molécula *ectoM* WNV en columna de inmunoafinidad.

- El plásmido TRIPdeltaU3CMV[95-114]EGFP[215-255]WNV se cotransfecta en células 293T con los plásmidos 8.7 y G-VSV para producir partículas virales pseudotipadas por la envoltura G del virus de estomatitis vesicular (VSV), que contiene las proteínas internas del virus de la inmunodeficiencia adquirida (VIH) y moléculas de ARN quiméricas CMV[95-114]EGFP[215-255]WNV. La infección de las células diana por el vector recombinado no replicativo permite la integración en el genoma celular del ADN CMV[95-114]EGFP[215-255]WNV y la expresión estable del ectodominio de la proteína M-WN bajo el control del promotor CMV.

3- La proteína NS 1

- 45 El ARN genómico extraído de los viriones IS-98-ST1 purificados se amplifica mediante la técnica RT-PCR (kit Titan One Tube RT-PCR; Roche Biochemicals #1939 823) con ayuda del siguiente par de cebadores:

- 5'TGGATGGGATCCAATATGCGTGATAGGTCC 3' (SEQ ID N° 9) que contiene el sitio de restricción *BamHI*, y

- 50 - 3'AAAAGGGTCAATGGTACCAGCATTTTAAGCATTCACGTT 3' (SEQ ID N° 10) que contiene el sitio de restricción *KpnI*.

- El ADNc que codifica la glucoproteína NS1 con su péptido señal (aminoácidos 767 a 1143 de la poliproteína viral) se clona entre los sitios *BamHI* y *KpnI* del vector retroviral TRIPdeltaU3 para producir el plásmido recombinado TRIPdeltaU3-CMV-NS1-WN (depositado el 9 de enero de 2002 en la *Collection Nationale de Culture de Microorganismes* del Instituto Pasteur de París, 28 rue du Docteur ROUX, 75724, PARIS Cedex 15 con el número I-2770). El plásmido TRIPdeltaU3-CMV-NS1-WN se cotransfecta en células 293T con los plásmidos 8.7 y G-VSV para producir partículas virales pseudotipadas por la envoltura G del virus de estomatitis vesicular (VSV), que contiene las

proteínas internas del virus de la inmunodeficiencia adquirida (VIH) y moléculas de ARN quiméricas CMV-NS1-WN. La infección de las células diana por el vector recombinado no replicativo permite la integración en el genoma celular del ADN CMV-NS1-WN y la expresión estable de la proteína NS1 del virus WN bajo el control del promotor CMV. La proteína NS1 de la cepa IS-98-ST1 del virus WN producida de este modo se usa para estudios estructurales, para la búsqueda de socios celulares de esta proteína en columna de inmutafinidad, y para la producción de anticuerpos monoespecíficos en conejo.

Ejemplo 4: Los ratones silvestres y de líneas consanguíneas de laboratorio se diferencian por su sensibilidad a la infección por la cepa neuroinvasiva IS-98-ST1 del virus West-Nile.

1- Las líneas de ratón y las células sensibles

a) Líneas de ratón sensibles

A ratones de líneas consanguíneas sensibles *Flv^s* (BALB/c) de 6 semanas de edad se les inoculan por vía intraperitoneal 100 UFF de la cepa IS-98-ST1 del virus West-Nile (UFF: DL50 = 10), preparada como se describe en el ejemplo 1.

El 100% de estos ratones muere en un tiempo medio de mortalidad de 9 ± 2 días (Figura 2).

La cinética de propagación de la cepa IS-98ST1 en el sistema nervioso central del ratón sensible (BALB/c) se analizó a partir de extractos de cerebro de ratones infectados valorados cuantitativamente en células AP61, de acuerdo con la técnica descrita en Després y col. (J. Virol., 1998, 72, 823-829). Los resultados muestran que el virus se detecta en el sistema nervioso central (SNC) murino al 5º día de la infección y que la producción viral es máxima al 7º día (Figura 3). Al 9º día de la infección, el virus ya no se detecta en el SNC murino (Figura 3).

La replicación del virus WN en el SNC y los órganos periféricos de los ratones infectados por la cepa IS-98-ST1 también se detecta mediante inmunohistología, de acuerdo con los protocolos clásicos tales como los descritos en Després y col., 1998 (mencionado anteriormente) y mediante hibridación *in situ*, de acuerdo con los protocolos descritos en el ejemplo 3.

Los anticuerpos séricos específicamente dirigidos contra las proteínas del virus WN se valoran cuantitativamente mediante ELISA de acuerdo con el protocolo descrito en Després y col., 1993 (mencionado anteriormente), usando la cepa IS-98-ST1 purificada en gradiente de sacarosa tal como se describe en el ejemplo 1 como antígeno. Los resultados muestran que los anticuerpos séricos aparecen al 5º día de la infección y se detectan significativamente al 7º día (figura 2).

b) Células sensibles

b1) Cultivos primarios

Neuronas primarias y astrocitos del SNC de ratones sensibles homocigóticos para el alelo *Flv^s* (ratón Swiss, Enero) se preparan de acuerdo con los protocolos clásicos. Las células son infectadas por la cepa IS-98-ST1 a una multiplicidad de infección de 20 UFF por célula (m.i. de 20). El efecto citopático se observa al microscopio óptico, la producción viral se analiza por valoración cuantitativa en células AP61 como se ha descrito anteriormente en el ejemplo 1 y la expresión de los antígenos virales se analiza por radioinmunoprecipitación con ayuda de un suero inmune de ratón anti-West-Nile, de acuerdo con los protocolos clásicos tales como los descritos en Duarte Dos Santos y col. (Virology, 2000. 274: 292-308).

Los resultados muestran que el 80% de las neuronas en cultivo producen los antígenos virales:

- su perfil en gel de poliacrilamida-SDS se presenta en la figura 4A,

- la producción viral es de $[3,0 \pm 1,5] \times 10^6$ UFF/ml después de 20 h de infección y de $[7,0 \pm 0,5] \times 10^7$ UFF/ml a 40 h,

- los efectos citopáticos (ECP) de tipo necrótico se observan después de 48 h de infección vírica.

Por el contrario, los astrocitos del SNC murino no son permisivos a la replicación del virus WN cepa IS-98-ST1.

b2) Líneas celulares

Células de neuroblastoma murino Neuro 2a y células de hepatoma humano HepG2, cultivadas en condiciones clásicas tal como se describen en Marianneau y col. (J. Virol., 1996, 77: 2547-2554) se infectan a diferente multiplicidad de infección por el virus WN cepa IS-98-ST1, preparado como se describe en el ejemplo 1. El efecto citopático se observa al microscopio óptico, la producción viral se analiza mediante valoración cuantitativa en células AP61 como se ha descrito anteriormente en el ejemplo 1 y la expresión de los antígenos virales se analiza por radioinmunoprecipitación con ayuda de un suero inmune de ratón anti-West-Nile, de acuerdo con los protocolos clásicos tales como los descritos en Duarte Dos Santos y col., Virol., 2000, 274, 292-308.

Los resultados muestran que las células de neuroblastoma murino Neuro 2a son permisivas a la replicación de la cepa IS-98-ST1 del virus WN. Es necesaria una m.i. de 4 para infectar el 80% de las células Neuro 2a en monocapa. La producción viral es de 10^7 UFF/ml (m.i. de 4) después de 40 h de infección y la muerte celular por necrosis es masiva (Figura 5). La cinética de producción de los principales antígenos prM, E y NS1 a partir de la poliproteína viral presentada en la figura 4B muestra que el tiempo necesario para la formación de la mitad de la glicoproteína de la envoltura E es de aproximadamente 30 min. La proteína E de la cepa IS-98-ST1 parece poseer solamente un único resto N-glicano (figura 4C).

Los resultados muestran también que las células de hepatoma humano HepG2 son permisivas a la replicación de la cepa IS-98-ST1 del virus M/N. A una m.i. de 10, la producción viral es de $[2 \pm 1] \times 10^6$ UFF/ml después de 48 h de infección y los ECP se observan a partir de 72 h.

2- Los ratones resistentes

Las líneas de ratones resistentes (*Flv*) que derivan de ratones silvestres de la especie *Mus spretus* (SEG/Pas y STF/Pas), *Mus musculus musculus* (MBT/Pas, MAI/pas) y *Mus musculus domesticus* (WMP/Pas), se inoculan por vía intraperitoneal con 1000 UFF (100 DL50) de la cepa IS-98-ST1 preparada de acuerdo con el protocolo descrito en el ejemplo 1.

Al contrario que los ratones de laboratorio, que son sensibles a la infección por la cepa IS-98-ST1 y mueren en diez días, estos ratones que derivan de ratones silvestres son resistentes a la inoculación de la cepa IS-98-ST1 y, sin embargo, permisivos a la replicación de la cepa IS-98-ST1. En efecto, la infección vírica de ratones que derivan de ratones silvestres es asintomática aunque el virus se multiplica *in toto* como demuestra la producción de anticuerpos séricos anti-WN a valoraciones cuantitativas altas; en ELISA, las valoraciones cuantitativas de los sueros a la dilución 1: 100 para 10^6 UFF de virión purificado IS-98-ST1 son superiores a 1 unidad de D.O. a 450 nm.

Los ratones resistentes a la infección vírica se usan para la producción de sueros inmunes específicamente dirigidos contra las proteínas de la cepa IS-98-ST1 del virus WN. Tres semanas después de la inoculación del virus WN, los sueros extraídos de ratones resistentes (0,045 ml por ratón) se mezclan, se descomplementan durante 30 min a 56°C y después se diluyen a 1: 10 en DPBS* (V/V) suplementado con el 0,2% (V/V) de Albúmina de suero bovino (Life Technologies) y el 0,05% (P/V) de azida sódica. Los sueros diluidos se distribuyen en 0,2 ml y se conservan a -20°C. Los sueros inmunes dirigidos contra la cepa IS-98-ST1 se usan a las diluciones finales de 1: 500 para la inmunofluorescencia indirecta y a 1: 1000 para la inmunoprecipitación de las proteínas virales radiomarcadas.

Ejemplo 5: Uso de la cepa IS-98-ST1 del virus West-Nile para identificar los genes celulares implicados en la sensibilidad del hospedador a la infección por los virus de la familia de los *Flaviviridae*.

1) Procedimientos

a) Modelo de análisis de la resistencia a la infección por *Flaviviridae* (figura 6)

Se cruzan ratones macho de las líneas resistentes MAI/Pas y MBT/Pas con ratones hembra de las líneas sensibles C57BL/6 y BALB/c. Los ratones macho de la generación F1 se retrocruzan con ratones hembra de las líneas sensibles C57BL/6 y BALB/c para dar una generación de ratones de primer retrocruzamiento (BC1).

A ratones BC1 de 5 semanas de edad se les inocula por vía intraperitoneal la cepa IS-98-ST1, preparada de acuerdo con el protocolo descrito en el ejemplo 1, en las condiciones descritas en el ejemplo 2.

Los animales se observan todos los días y las tasas de mortalidad y de supervivencia se determinan 14 días

después de la infección.

b) Genotipado de los alelos *Flv*

- 5 Los alelos *Flv* de los individuos BC1 se mapearon mediante PCR genómica con ayuda de cebadores específicos de 16 microsatélites del cromosoma 5 (Catalogue Research Genetics) que rodean al locus *Flv* (figuras 7-9), de acuerdo con las técnicas habituales de biología molecular usando los protocolos convencionales tales como los descritos en Current Prolocols in Molecular Biology (Frederick M. AUSUBEL, 2000, Wiley and son Inc, Library of Congress, EE.UU.).

10

2) Resultados

- El análisis de la distribución de los alelos *Flv* en los ratones BC1 sensibles y resistentes a la infección por la cepa IS-98ST1 muestra que un alelo *Flv*^r es suficiente para otorgar la resistencia a la infección (figura 10). Los resultados
15 muestran también que en este modelo existe una correlación perfecta entre el fenotipo resistente y la presencia del alelo *Flv*^r y una correlación casi perfecta entre el fenotipo sensible y la ausencia del alelo *Flv*^r (figura 10).

El genotipado de los alelos *Flv* muestra que el locus *Flv* se localiza en una región de 0,2 cM que contiene el gen OAS 1 (figuras 7-9).

20

LISTA DE SECUENCIAS

- < 110 > INSTITUT PASTEUR
25 KIMRON VETERINARY INSTITUTE
DESPRES Philippe
DEUBEL Vincent
GUENET Jean-Louis
DROUET Marie-Thérèse
30 MALKINSON Mertyn
BANET Caroline
FRENKIEL Marie-Pascale
COURAGEOT Marie-Pierre
COULILABY Fasséli
35 CATTEAU Adeline
FLAMAND Marie
WEBER Patrick
CECCALDI Pierre-Emmanuel

- 40 < 120 > CEPA NEUROVIRULENTA DEL VIRUS WEST NILE Y SUS APLICACIONES

< 130 > 226CAS93EXT

< 140 >

- 45 < 141 >

< 160 > 11

< 170 > PatentIn Ver. 2.1

50

< 210 > 1

< 211 > 11029

< 212 > ARN

< 213 > Flavivirus sp.

55

< 220 >

< 221 > CDS

< 222 > (97)..(10395)

< 400 > 1

```

agtagttcgc ctgtgtgagc tgacaaactt agtagtgttt gtgaggatta acaacaatta 60
acacagtgcg agctgtttct tagcacgaag atctcg atg tct aag aaa cca gga 114
                               Met Ser Lys Lys Pro Gly
                               1                               5
ggg ccc ggc aag agc cgg gct gtc aat atg cta aaa cgc gga atg ccc 162
Gly Pro Gly Lys Ser Arg Ala Val Asn Met Leu Lys Arg Gly Met Pro
                               10                               15                               20
cgc gtg ttg tcc ttg att gga ctg aag agg gct atg ttg agc ctg atc 210
Arg Val Leu Ser Leu Ile Gly Leu Lys Arg Ala Met Leu Ser Leu Ile
                               25                               30                               35
gac ggc aag ggg cca ata cga ttt gtg ttg gct ctc ttg gcg ttc ttc 258
Asp Gly Lys Gly Pro Ile Arg Phe Val Leu Ala Leu Leu Ala Phe Phe
                               40                               45                               50
agg ttc aca gca att gct ccg acc cga gca gtg ctg gat cga tgg aga 306
Arg Phe Thr Ala Ile Ala Pro Thr Arg Ala Val Leu Asp Arg Trp Arg
                               55                               60                               65                               70

```

ggt gtg aat aaa caa aca gcg atg aaa cac ctt ctg agt ttt aag aag	354
Gly Val Asn Lys Gln Thr Ala Met Lys His Leu Leu Ser Phe Lys Lys	
75 80 85	
gaa cta ggg acc ttg acc agt gct atc aat cgg cgg agc tca aaa caa	402
Glu Leu Gly Thr Leu Thr Ser Ala Ile Asn Arg Arg Ser Ser Lys Gln	
90 95 100	
aag aaa aga gga gga aag acc gga att gca gtc atg att ggc ctg atc	450
Lys Lys Arg Gly Gly Lys Thr Gly Ile Ala Val Met Ile Gly Leu Ile	
105 110 115	
gcc agc gta gga gca gtt acc ctc tct aac ttc caa ggg aag gtg atg	498
Ala Ser Val Gly Ala Val Thr Leu Ser Asn Phe Gln Gly Lys Val Met	
120 125 130	
atg acg gta aat gct act gac gtc aca gat gtc atc acg att cca aca	546
Met Thr Val Asn Ala Thr Asp Val Thr Asp Val Ile Thr Ile Pro Thr	
135 140 145 150	
gct gct gga aag aac cta tgc att gtc aga gca atg gat gtg gga tac	594
Ala Ala Gly Lys Asn Leu Cys Ile Val Arg Ala Met Asp Val Gly Tyr	
155 160 165	
atg tgc gat gat act atc act tat gaa tgc cca gtg ctg tgc gct ggt	642
Met Cys Asp Asp Thr Ile Thr Tyr Glu Cys Pro Val Leu Ser Ala Gly	
170 175 180	
aat gat cca gaa gac atc gac tgt tgg tgc aca aag tca gca gtc tac	690
Asn Asp Pro Glu Asp Ile Asp Cys Trp Cys Thr Lys Ser Ala Val Tyr	
185 190 195	
gtc agg tat gga aga tgc acc aag aca cgc cac tca aga cgc agt cgg	738
Val Arg Tyr Gly Arg Cys Thr Lys Thr Arg His Ser Arg Arg Ser Arg	
200 205 210	
agg tca ctg aca gtg cag aca cac gga gaa agc act cta gcg aac aag	786
Arg Ser Leu Thr Val Gln Thr His Gly Glu Ser Thr Leu Ala Asn Lys	
215 220 225 230	
aag ggg gct tgg atg gac agc acc aag gcc aca agg tat ttg gta aaa	834
Lys Gly Ala Trp Met Asp Ser Thr Lys Ala Thr Arg Tyr Leu Val Lys	
235 240 245	
aca gaa tca tgg atc ttg agg aac cct gga tat gcc ctg gtg gca gcc	882
Thr Glu Ser Trp Ile Leu Arg Asn Pro Gly Tyr Ala Leu Val Ala Ala	
250 255 260	
gtc att ggt tgg atg ctt ggg agc aac acc atg cag aga gtt gtg ttt	930
Val Ile Gly Trp Met Leu Gly Ser Asn Thr Met Gln Arg Val Val Phe	
265 270 275	
gtc gtg cta ttg ctt ttg gtg gcc cca gct tac agc ttt aac tgc ctt	978
Val Val Leu Leu Leu Leu Val Ala Pro Ala Tyr Ser Phe Asn Cys Leu	
280 285 290	
gga atg agc aac aga gac ttc ttg gaa gga gtg tct gga gca aca tgg	1026
Gly Met Ser Asn Arg Asp Phe Leu Glu Gly Val Ser Gly Ala Thr Trp	
295 300 305 310	

gtg gat ttg gtt ctc gaa ggc gac agc tgc gtg act atc atg tct aag Val Asp Leu Val Leu Glu Gly Asp Ser Cys Val Thr Ile Met Ser Lys 315 320 325	1074
gac aag cct acc atc gat gtg aag atg atg aat atg gag gcg gcc aac Asp Lys Pro Thr Ile Asp Val Lys Met Met Asn Met Glu Ala Ala Asn 330 335 340	1122
ctg gca gag gtc cgc agt tat tgc tat ttg gct acc gtc agc gat ctc Leu Ala Glu Val Arg Ser Tyr Cys Tyr Leu Ala Thr Val Ser Asp Leu 345 350 355	1170
tcc acc aaa gct gcg tgc ccg acc atg gga gaa gct cac aat gac aaa Ser Thr Lys Ala Ala Cys Pro Thr Met Gly Glu Ala His Asn Asp Lys 360 365 370	1218
cgt gct gac cca gct ttt gtg tgc aga caa gga gtg gtg gac agg ggc Arg Ala Asp Pro Ala Phe Val Cys Arg Gln Gly Val Val Asp Arg Gly 375 380 385 390	1266
tgg ggc aac ggc tgc gga cta ttt ggc aaa gga agc att gac aca tgc Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys Gly Ser Ile Asp Thr Cys 395 400 405	1314
gcc aaa ttt gcc tgc tct acc aag gca ata gga aga acc atc ttg aaa Ala Lys Phe Ala Cys Ser Thr Lys Ala Ile Gly Arg Thr Ile Leu Lys 410 415 420	1362
gag aat atc aag tac gaa gtg gcc att ttt gtc cat gga cca act act Glu Asn Ile Lys Tyr Glu Val Ala Ile Phe Val His Gly Pro Thr Thr 425 430 435	1410
gtg gag tgc cac gga aac tac tcc aca cag gtt gga gcc act cag gca Val Glu Ser His Gly Asn Tyr Ser Thr Gln Val Gly Ala Thr Gln Ala 440 445 450	1458
ggg aga ttc agc atc act cct gcg gcg cct tca tac aca cta aag ctt Gly Arg Phe Ser Ile Thr Pro Ala Ala Pro Ser Tyr Thr Leu Lys Leu 455 460 465 470	1506
gga gaa tat gga gag gtg aca gtg gac tgt gaa cca cgg tca ggg att Gly Glu Tyr Gly Glu Val Thr Val Asp Cys Glu Pro Arg Ser Gly Ile 475 480 485	1554
gac acc aat gca tac tac gtg atg act gtt gga aca aag acg ttc ttg Asp Thr Asn Ala Tyr Tyr Val Met Thr Val Gly Thr Lys Thr Phe Leu 490 495 500	1602
gtc cat cgt gag tgg ttc atg gac ctc aac ctc cct tgg agc agt gct Val His Arg Glu Trp Phe Met Asp Leu Asn Leu Pro Trp Ser Ser Ala 505 510 515	1650
gga agt act gtg tgg agg aac aga gag acg tta atg gag ttt gag gaa Gly Ser Thr Val Trp Arg Asn Arg Glu Thr Leu Met Glu Phe Glu Glu 520 525 530	1698
cca cac gcc acg aag cag tct gtg ata gca ttg ggc tca caa gag gga Pro His Ala Thr Lys Gln Ser Val Ile Ala Leu Gly Ser Gln Glu Gly 535 540 545 550	1746

gct ctg cat caa gct ttg gct gga gcc att cct gtg gaa ttt tca agc	1794
Ala Leu His Gln Ala Leu Ala Gly Ala Ile Pro Val Glu Phe Ser Ser	
555 560 565	
aac act gtc aag ttg acg tcg ggt cat ttg aag tgt aga gtg aag atg	1842
Asn Thr Val Lys Leu Thr Ser Gly His Leu Lys Cys Arg Val Lys Met	
570 575 580	
gaa aaa ttg cag ttg aag gga aca acc tat ggc gtc tgt tca aag gct	1890
Glu Lys Leu Gln Leu Lys Gly Thr Thr Tyr Gly Val Cys Ser Lys Ala	
585 590 595	
ttc aag ttt ctt ggg act ccc gca gac aca ggt cac ggc act gtg gtg	1938
Phe Lys Phe Leu Gly Thr Pro Ala Asp Thr Gly His Gly Thr Val Val	
600 605 610	
ttg gaa ttg cag tac act ggc acg gat gga cct tgc aaa gtt cct atc	1986
Leu Glu Leu Gln Tyr Thr Gly Thr Asp Gly Pro Cys Lys Val Pro Ile	
615 620 625 630	
tcg tca gtg gct tca ttg aac gac cta acg cca gtg ggc aga ttg gtc	2034
Ser Ser Val Ala Ser Leu Asn Asp Leu Thr Pro Val Gly Arg Leu Val	
635 640 645	
act gtc aac cct ttt gtt tca gtg gcc acg gcc aac gct aag gtc ctg	2082
Thr Val Asn Pro Phe Val Ser Val Ala Thr Ala Asn Ala Lys Val Leu	
650 655 660	
att gaa ttg gaa cca ccc ttt gga gac tca tac ata gtg gtg ggc aga	2130
Ile Glu Leu Glu Pro Pro Phe Gly Asp Ser Tyr Ile Val Val Gly Arg	
665 670 675	
gga gaa caa cag att aat cac cat tgg cac aag tct gga agc agc att	2178
Gly Glu Gln Gln Ile Asn His His Trp His Lys Ser Gly Ser Ser Ile	
680 685 690	
ggc aaa gcc ttt aca acc acc ctc aaa gga gcg cag aga cta gcc gct	2226
Gly Lys Ala Phe Thr Thr Thr Leu Lys Gly Ala Gln Arg Leu Ala Ala	
695 700 705 710	
cta gga gac aca gct tgg gac ttt gga tca gtt gga ggg gtg ttc acc	2274
Leu Gly Asp Thr Ala Trp Asp Phe Gly Ser Val Gly Gly Val Phe Thr	
715 720 725	
tca gtt ggg aag gct gtc cat caa gtg ttc gga gga gca ttc cgc tca	2322
Ser Val Gly Lys Ala Val His Gln Val Phe Gly Gly Ala Phe Arg Ser	
730 735 740	
ctg ttc gga ggc atg tcc tgg ata acg caa gga ttg ctg ggg gct ctc	2370
Leu Phe Gly Gly Met Ser Trp Ile Thr Gln Gly Leu Leu Gly Ala Leu	
745 750 755	
ctg ttg tgg atg ggc atc aat gct cgt gat agg tcc ata gct ctc acg	2418
Leu Leu Trp Met Gly Ile Asn Ala Arg Asp Arg Ser Ile Ala Leu Thr	
760 765 770	
ttt ctc gca gtt gga gga gtt ctg ctc ttc ctc tcc gtg aac gtg cac	2466
Phe Leu Ala Val Gly Gly Val Leu Leu Phe Leu Ser Val Asn Val His	
775 780 785 790	

gct gac act ggg tgt gcc ata gac atc agc cgg caa gag ctg aga tgt Ala Asp Thr Gly Cys Ala Ile Asp Ile Ser Arg Gln Glu Leu Arg Cys 795 800 805	2514
gga aat gga gtg ttc ata cac aat gat gtg gag gct tgg atg gac cgg Gly Asn Gly Val Phe Ile His Asn Asp Val Glu Ala Trp Met Asp Arg 810 815 820	2562
tac aag tat tac cct gaa acg cca caa ggc cta gcc aag atc att cag Tyr Lys Tyr Tyr Pro Glu Thr Pro Gln Gly Leu Ala Lys Ile Ile Gln 825 830 835	2610
aaa gct cat aag gaa gga gtg tgc ggt cta cga tca gtt tcc aga ctg Lys Ala His Lys Glu Gly Val Cys Gly Leu Arg Ser Val Ser Arg Leu 840 845 850	2658
gag cat caa atg tgg gaa gca gtg aag gac gag ctg aac act ctt ttg Glu His Gln Met Trp Glu Ala Val Lys Asp Glu Leu Asn Thr Leu Leu 855 860 865 870	2706
aag gag aat ggt gtg gac ctt agt gtc gtg gtt gag aaa cag gag gga Lys Glu Asn Gly Val Asp Leu Ser Val Val Val Glu Lys Gln Glu Gly 875 880 885	2754
atg tac aag tca gca cct aaa cgc ctc acc gcc acc acg gaa aaa ttg Met Tyr Lys Ser Ala Pro Lys Arg Leu Thr Ala Thr Thr Glu Lys Leu 890 895 900	2802
gaa att ggc tgg aag gcc tgg gga aag agt att tta ttt gca cca gaa Glu Ile Gly Trp Lys Ala Trp Gly Lys Ser Ile Leu Phe Ala Pro Glu 905 910 915	2850
ctc gcc aac aac acc ttt gtg gtt gat ggt ccg gag acc aag gaa tgt Leu Ala Asn Asn Thr Phe Val Val Asp Gly Pro Glu Thr Lys Glu Cys 920 925 930	2898
ccg act cag aat cgc gct tgg aat agc tta gaa gtg gag gat ttt gga Pro Thr Gln Asn Arg Ala Trp Asn Ser Leu Glu Val Glu Asp Phe Gly 935 940 945 950	2946
ttt ggt ctc acc agc act cgg atg ttc ctg aag gtc aga gag agc aac Phe Gly Leu Thr Ser Thr Arg Met Phe Leu Lys Val Arg Glu Ser Asn 955 960 965	2994
aca act gaa tgt gac tcg aag atc att gga acg gct gtc aag aac aac Thr Thr Glu Cys Asp Ser Lys Ile Ile Gly Thr Ala Val Lys Asn Asn 970 975 980	3042
ttg gcg atc cac agt gac ctg tcc tat tgg att gaa agc agg ctc aat Leu Ala Ile His Ser Asp Leu Ser Tyr Trp Ile Glu Ser Arg Leu Asn 985 990 995	3090
gat acg tgg aag ctt gaa agg gca gtt ctg ggt gaa gtc aaa tca tgt Asp Thr Trp Lys Leu Glu Arg Ala Val Leu Gly Glu Val Lys Ser Cys 1000 1005 1010	3138
acg tgg cct gag acg cat acc ttg tgg ggc gat gga atc ctt gag agt Thr Trp Pro Glu Thr His Thr Leu Trp Gly Asp Gly Ile Leu Glu Ser 1015 1020 1025 1030	3186

gac ttg ata ata cca gtc aca ctg gcg gga cca cga agc aat cac aat	3234
Asp Leu Ile Ile Pro Val Thr Leu Ala Gly Pro Arg Ser Asn His Asn	
1035 1040 1045	
cgg aga cct ggg tac aag aca caa aac cag ggc cca tgg gac gaa ggc	3282
Arg Arg Pro Gly Tyr Lys Thr Gln Asn Gln Gly Pro Trp Asp Glu Gly	
1050 1055 1060	
cgg gta gag att gac ttc gat tac tgc cca gga act acg gtc acc ctg	3330
Arg Val Glu Ile Asp Phe Asp Tyr Cys Pro Gly Thr Thr Val Thr Leu	
1065 1070 1075	
agt gag agc tgc gga cac cgt gga cct gcc act cgc acc acc aca gag	3378
Ser Glu Ser Cys Gly His Arg Gly Pro Ala Thr Arg Thr Thr Thr Glu	
1080 1085 1090	
agc gga aag ttg ata aca gat tgg tgc tgc agg agc tgc acc tta cca	3426
Ser Gly Lys Leu Ile Thr Asp Trp Cys Cys Arg Ser Cys Thr Leu Pro	
1095 1100 1105 1110	
cca ctg cgc tac caa act gac agc ggc tgt tgg tat ggt atg gag atc	3474
Pro Leu Arg Tyr Gln Thr Asp Ser Gly Cys Trp Tyr Gly Met Glu Ile	
1115 1120 1125	
aga cca cag aga cat gat gaa aag acc ctc gtg cag tca caa gtg aat	3522
Arg Pro Gln Arg His Asp Glu Lys Thr Leu Val Gln Ser Gln Val Asn	
1130 1135 1140	
gct tat aat gct gat atg att gac cct ttt cag ttg ggc ctt ctg gtc	3570
Ala Tyr Asn Ala Asp Met Ile Asp Pro Phe Gln Leu Gly Leu Leu Val	
1145 1150 1155	
gtg ttc ttg gcc acc cag gag gtc ctt cgc aag agg tgg aca gcc aag	3618
Val Phe Leu Ala Thr Gln Glu Val Leu Arg Lys Arg Trp Thr Ala Lys	
1160 1165 1170	
atc agc atg cca gct ata ctg att gct ctg cta gtc ctg gtg ttt ggg	3666
Ile Ser Met Pro Ala Ile Leu Ile Ala Leu Leu Val Leu Val Phe Gly	
1175 1180 1185 1190	
ggc att act tac act gat gtg tta cgc tat gtc atc ttg gtg ggc gca	3714
Gly Ile Thr Tyr Thr Asp Val Leu Arg Tyr Val Ile Leu Val Gly Ala	
1195 1200 1205	
gct ttc gca gaa tct aat tcg gga gga gac gtg gta cac ttg gcg ctc	3762
Ala Phe Ala Glu Ser Asn Ser Gly Gly Asp Val Val His Leu Ala Leu	
1210 1215 1220	
atg gcg acc ttc aag ata caa cca gtg ttt atg gtg gca tcg ttt ctc	3810
Met Ala Thr Phe Lys Ile Gln Pro Val Phe Met Val Ala Ser Phe Leu	
1225 1230 1235	
aaa gcg aga tgg acc aac cag gag aac att ttg ttg atg ttg gcg gct	3858
Lys Ala Arg Trp Thr Asn Gln Glu Asn Ile Leu Leu Met Leu Ala Ala	
1240 1245 1250	
gtt ttc ttt caa atg gct tat cac gat gcc cgc caa att ctg ctc tgg	3906
Val Phe Phe Gln Met Ala Tyr His Asp Ala Arg Gln Ile Leu Leu Trp	
1255 1260 1265 1270	

gag atc cct gat gtg ttg aat tca ctg gcg gta gct tgg atg ata ctg Glu Ile Pro Asp Val Leu Asn Ser Leu Ala Val Ala Trp Met Ile Leu 1275 1280 1285	3954
aga gcc ata aca ttc aca acg aca tca aat gtg gtc gtc ccg ctg cta Arg Ala Ile Thr Phe Thr Thr Thr Ser Asn Val Val Val Pro Leu Leu 1290 1295 1300	4002
gcc ctg cta aca ccc cgg ctg aga tgc ttg aat ctg gat gtg tac agg Ala Leu Leu Thr Pro Arg Leu Arg Cys Leu Asn Leu Asp Val Tyr Arg 1305 1310 1315	4050
ata ctg ctg ttg atg gtc gga ata ggc agc ttg atc agg gag aag agg Ile Leu Leu Leu Met Val Gly Ile Gly Ser Leu Ile Arg Glu Lys Arg 1320 1325 1330	4098
agt gca gct gca aaa aag aaa gga gca agt ctg cta tgc ttg gct cta Ser Ala Ala Ala Lys Lys Lys Gly Ala Ser Leu Leu Cys Leu Ala Leu 1335 1340 1345 1350	4146
gcc tca aca gga ctt ttc aac ccc atg atc ctt gct gct gga ctg att Ala Ser Thr Gly Leu Phe Asn Pro Met Ile Leu Ala Ala Gly Leu Ile 1355 1360 1365	4194
gca tgt gat ccc aac cgt aaa cgc gga tgg ccc gca act gaa gtg atg Ala Cys Asp Pro Asn Arg Lys Arg Gly Trp Pro Ala Thr Glu Val Met 1370 1375 1380	4242
aca gct gtc ggc cta atg ttt gcc atc gtc gga ggg ctg gca gag ctt Thr Ala Val Gly Leu Met Phe Ala Ile Val Gly Gly Leu Ala Glu Leu 1385 1390 1395	4290
gac att gac tcc atg gcc att cca atg act atc gcg ggg ctc atg ttt Asp Ile Asp Ser Met Ala Ile Pro Met Thr Ile Ala Gly Leu Met Phe 1400 1405 1410	4338
gct gct ttc gtg att tct ggg aaa tca aca gat atg tgg att gag aga Ala Ala Phe Val Ile Ser Gly Lys Ser Thr Asp Met Trp Ile Glu Arg 1415 1420 1425 1430	4386
acg gcg gac att tcc tgg gaa agt gat gca gaa att aca ggc tcg agc Thr Ala Asp Ile Ser Trp Glu Ser Asp Ala Glu Ile Thr Gly Ser Ser 1435 1440 1445	4434
gaa aga gtt gat gtt cgg ctt gat gat ggt gaa aac ttc cag ctc atg Glu Arg Val Asp Val Arg Leu Asp Asp Gly Glu Asn Phe Gln Leu Met 1450 1455 1460	4482
aat gat cca gga gca cct tgg aag ata tgg atg ctc aga atg gtc tgt Asn Asp Pro Gly Ala Pro Trp Lys Ile Trp Met Leu Arg Met Val Cys 1465 1470 1475	4530
ctc gcg att agt gcg tac acc ccc tgg gca atc ttg ccc tca gta gtt Leu Ala Ile Ser Ala Tyr Thr Pro Trp Ala Ile Leu Pro Ser Val Val 1480 1485 1490	4578
gga ttt tgg ata act ctc caa tac aca aag aga gga ggt gtg ttg tgg Gly Phe Trp Ile Thr Leu Gln Tyr Thr Lys Arg Gly Gly Val Leu Trp 1495 1500 1505 1510	4626

gac act ccc tca cca aag gag tac aaa aag ggg gac acg acc acc ggc	4674
Asp Thr Pro Ser Pro Lys Glu Tyr Lys Lys Gly Asp Thr Thr Thr Gly	
1515 1520 1525	
gtc tac agg atc atg act cgt ggg ctg ctc ggc agt tat caa gca gga	4722
Val Tyr Arg Ile Met Thr Arg Gly Leu Leu Gly Ser Tyr Gln Ala Gly	
1530 1535 1540	
gcg ggc gtg atg gtt gaa ggt gtt ttc cac acc ctt tgg cat aca aca	4770
Ala Gly Val Met Val Glu Gly Val Phe His Thr Leu Trp His Thr Thr	
1545 1550 1555	
aaa gga gcc gct ttg atg agc gga gag ggc cgc ctg gac cca tac tgg	4818
Lys Gly Ala Ala Leu Met Ser Gly Glu Gly Arg Leu Asp Pro Tyr Trp	
1560 1565 1570	
ggc agt gtc aag gag gat cga ctt tgt tac gga gga ccc tgg aaa ttg	4866
Gly Ser Val Lys Glu Asp Arg Leu Cys Tyr Gly Gly Pro Trp Lys Leu	
1575 1580 1585 1590	
cag cac aag tgg aac ggg cag gat gag gtg cag atg att gtg gtg gaa	4914
Gln His Lys Trp Asn Gly Gln Asp Glu Val Gln Met Ile Val Val Glu	
1595 1600 1605	
cct ggc aag aac gtt aag aac gtc cag acg aaa cca ggg gtg ttc aaa	4962
Pro Gly Lys Asn Val Lys Asn Val Gln Thr Lys Pro Gly Val Phe Lys	
1610 1615 1620	
aca cct gaa gga gaa atc ggg gcc gtg act ttg gac ttc ccc act gga	5010
Thr Pro Glu Gly Glu Ile Gly Ala Val Thr Leu Asp Phe Pro Thr Gly	
1625 1630 1635	
aca tca ggc tca cca ata gtg gac aaa aac ggt gat gtg att ggg ctt	5058
Thr Ser Gly Ser Pro Ile Val Asp Lys Asn Gly Asp Val Ile Gly Leu	
1640 1645 1650	
tat ggc aat gga gtc ata atg ccc aac ggc tca tac ata agc gcg ata	5106
Tyr Gly Asn Gly Val Ile Met Pro Asn Gly Ser Tyr Ile Ser Ala Ile	
1655 1660 1665 1670	
gtg cag ggt gaa agg atg gat gag cca atc cca gcc gga ttc gaa cct	5154
Val Gln Gly Glu Arg Met Asp Glu Pro Ile Pro Ala Gly Phe Glu Pro	
1675 1680 1685	
gag atg ctg agg aaa aaa cag atc act gta ctg gat ctc cat ccc ggc	5202
Glu Met Leu Arg Lys Lys Gln Ile Thr Val Leu Asp Leu His Pro Gly	
1690 1695 1700	
gcc ggt aaa aca agg agg att ctg cca cag atc atc aaa gag gcc ata	5250
Ala Gly Lys Thr Arg Arg Ile Leu Pro Gln Ile Ile Lys Glu Ala Ile	
1705 1710 1715	
aac aga aga ctg aga aca gcc gtg cta gca cca acc agg gtt gtg gct	5298
Asn Arg Arg Leu Arg Thr Ala Val Leu Ala Pro Thr Arg Val Val Ala	
1720 1725 1730	
gct gag atg gct gaa gca ctg aga gga ctg ccc atc cgg tac cag aca	5346
Ala Glu Met Ala Glu Ala Leu Arg Gly Leu Pro Ile Arg Tyr Gln Thr	
1735 1740 1745 1750	

tcc gca gtg ccc aga gaa cat aat gga aat gag att gtt gat gtc atg Ser Ala Val Pro Arg Glu His Asn Gly Asn Glu Ile Val Asp Val Met 1755 1760 1765	5394
tgt cat gct acc ctg acc cac agg ctg atg tct cct cac agg gtg ccg Cys His Ala Thr Leu Thr His Arg Leu Met Ser Pro His Arg Val Pro 1770 1775 1780	5442
aac tac aac ctg ttc gtg atg gat gag gct cat ttc acc gac cca gct Asn Tyr Asn Leu Phe Val Met Asp Glu Ala His Phe Thr Asp Pro Ala 1785 1790 1795	5490
agt atc gca gca aga ggt tac att tcc aca aag gtc gag cta ggg gag Ser Ile Ala Ala Arg Gly Tyr Ile Ser Thr Lys Val Glu Leu Gly Glu 1800 1805 1810	5538
gcg gcg gca ata ttc atg aca gcc acc cca cca ggc act tca gat cca Ala Ala Ala Ile Phe Met Thr Ala Thr Pro Pro Gly Thr Ser Asp Pro 1815 1820 1825 1830	5586
ttc cca gag tcc aat tca cca att tcc gac tta cag act gag atc ccg Phe Pro Glu Ser Asn Ser Pro Ile Ser Asp Leu Gln Thr Glu Ile Pro 1835 1840 1845	5634
gat cga gct tgg aac tct gga tac gaa tgg atc aca gaa tac acc ggg Asp Arg Ala Trp Asn Ser Gly Tyr Glu Trp Ile Thr Glu Tyr Thr Gly 1850 1855 1860	5682
aag acg gtt tgg ttt gtg cct agt gtc aag atg ggg aat gag att gcc Lys Thr Val Trp Phe Val Pro Ser Val Lys Met Gly Asn Glu Ile Ala 1865 1870 1875	5730
ctt tgc cta caa cgt gct gga aag aaa gta gtc caa ttg aac aga aag Leu Cys Leu Gln Arg Ala Gly Lys Lys Val Val Gln Leu Asn Arg Lys 1880 1885 1890	5778
tcg tac gag acg gag tac cca aaa tgt aag aac gat gat tgg gac ttt Ser Tyr Glu Thr Glu Tyr Pro Lys Cys Lys Asn Asp Asp Trp Asp Phe 1895 1900 1905 1910	5826
gtt atc aca aca gac ata tct gaa atg ggg gct aac ttc aag gcg agc Val Ile Thr Thr Asp Ile Ser Glu Met Gly Ala Asn Phe Lys Ala Ser 1915 1920 1925	5874
agg gtg att gac agc cgg aag agt gtg aaa cca acc atc ata aca gaa Arg Val Ile Asp Ser Arg Lys Ser Val Lys Pro Thr Ile Ile Thr Glu 1930 1935 1940	5922
gga gaa ggg aga gtg atc ctg gga gaa cca tct gca gtg aca gca gct Gly Glu Gly Arg Val Ile Leu Gly Glu Pro Ser Ala Val Thr Ala Ala 1945 1950 1955	5970
agt gcc gcc cag aga cgt gga cgt atc ggt aga aat ccg tcg caa gtt Ser Ala Ala Gln Arg Arg Gly Arg Ile Gly Arg Asn Pro Ser Gln Val 1960 1965 1970	6018
ggt gat gag tac tgt tat ggg ggg cac acg aat gaa gac gac tcg aac Gly Asp Glu Tyr Cys Tyr Gly Gly His Thr Asn Glu Asp Asp Ser Asn 1975 1980 1985 1990	6066

ttc gcc cat tgg act gag gca cga atc atg ccg gac aac atc aac atg Phe Ala His Trp Thr Glu Ala Arg Ile Met Pro Asp Asn Ile Asn Met 1995 2000 2005	6114
cca aac gga ctg atc gct caa ttc tac caa cca gag cgt gag aag gta Pro Asn Gly Leu Ile Ala Gln Phe Tyr Gln Pro Glu Arg Glu Lys Val 2010 2015 2020	6162
tat acc atg gag ggg gaa tac cgg ctc aga gga gaa gag agg aaa aac Tyr Thr Met Glu Gly Glu Tyr Arg Leu Arg Gly Glu Glu Arg Lys Asn 2025 2030 2035	6210
ttt ctg gaa ctg ttg agg act gca gat ctg cca gtt tgg ctg gct tac Phe Leu Glu Leu Leu Arg Thr Ala Asp Leu Pro Val Trp Leu Ala Tyr 2040 2045 2050	6258
aag gtt gca gcg gct gga gtg tca tac cac gac cgg agg tgg tgc ttt Lys Val Ala Ala Ala Gly Val Ser Tyr His Asp Arg Arg Trp Cys Phe 2055 2060 2065 2070	6306
gat ggt cct agg aca aac aca att tta gaa gac aac aac gaa gtg gaa Asp Gly Pro Arg Thr Asn Thr Ile Leu Glu Asp Asn Asn Glu Val Glu 2075 2080 2085	6354
gtc atc acg aag ctt ggt gaa agg aag att ctg agg ccg cgc tgg att Val Ile Thr Lys Leu Gly Glu Arg Lys Ile Leu Arg Pro Arg Trp Ile 2090 2095 2100	6402
gac gcc agg gtg tac tcg gat cac cag gca cta aag gcg ttc aag gac Asp Ala Arg Val Tyr Ser Asp His Gln Ala Leu Lys Ala Phe Lys Asp 2105 2110 2115	6450
ttc gcc tcg gga aaa cgt tct cag ata ggg ctc att gag gtt ctg gga Phe Ala Ser Gly Lys Arg Ser Gln Ile Gly Leu Ile Glu Val Leu Gly 2120 2125 2130	6498
aag atg cct gag cac ttc atg ggg aag aca tgg gaa gca ctt gac acc Lys Met Pro Glu His Phe Met Gly Lys Thr Trp Glu Ala Leu Asp Thr 2135 2140 2145 2150	6546
atg tac gtt gtg gcc act gca gag aaa gga gga aga gct cac aga atg Met Tyr Val Val Ala Thr Ala Glu Lys Gly Gly Arg Ala His Arg Met 2155 2160 2165	6594
gcc ctg gag gaa ctg cca gat gct ctt cag aca att gcc ttg att gcc Ala Leu Glu Glu Leu Pro Asp Ala Leu Gln Thr Ile Ala Leu Ile Ala 2170 2175 2180	6642
tta ttg agt gtg atg acc atg gga gta ttc ttc ctc ctc atg cag cgg Leu Leu Ser Val Met Thr Met Gly Val Phe Phe Leu Leu Met Gln Arg 2185 2190 2195	6690
aag ggc att gga aag ata ggt ttg gga ggc gct gtc ttg gga gtc gcg Lys Gly Ile Gly Lys Ile Gly Leu Gly Gly Ala Val Leu Gly Val Ala 2200 2205 2210	6738
acc ttt ttc tgt tgg atg gct gaa gtt cca gga acg aag atc gcc gga Thr Phe Phe Cys Trp Met Ala Glu Val Pro Gly Thr Lys Ile Ala Gly 2215 2220 2225 2230	6786

atg ttg ctg ctc tcc ctt ctc ttg atg att gtg cta att cct gag cca Met Leu Leu Leu Ser Leu Leu Leu Met Ile Val Leu Ile Pro Glu Pro 2235 2240 2245	6834
gag aag caa cgt tcg cag aca gac aac cag cta gcc gtg ttc ctg att Glu Lys Gln Arg Ser Gln Thr Asp Asn Gln Leu Ala Val Phe Leu Ile 2250 2255 2260	6882
tgt gtc atg acc ctt gtg agc gca gtg gca gcc aac gag atg ggt tgg Cys Val Met Thr Leu Val Ser Ala Val Ala Ala Asn Glu Met Gly Trp 2265 2270 2275	6930
cta gac aag acc aag agt gac ata agc agt ttg ttt ggg caa aga att Leu Asp Lys Thr Lys Ser Asp Ile Ser Ser Leu Phe Gly Gln Arg Ile 2280 2285 2290	6978
gag gtc aag gag aat ttc agc atg gga gag ttt ctt ctg gac ttg agg Glu Val Lys Glu Asn Phe Ser Met Gly Glu Phe Leu Leu Asp Leu Arg 2295 2300 2305 2310	7026
ccg gca aca gcc tgg tca ctg tac gct gtg aca aca gcg gtc ctc act Pro Ala Thr Ala Trp Ser Leu Tyr Ala Val Thr Thr Ala Val Leu Thr 2315 2320 2325	7074
cca ctg cta aag cat ttg atc acg tca gat tac atc aac acc tca ttg Pro Leu Leu Lys His Leu Ile Thr Ser Asp Tyr Ile Asn Thr Ser Leu 2330 2335 2340	7122
acc tca ata aac gtt cag gca agt gca cta ttc aca ctc gcg cga ggc Thr Ser Ile Asn Val Gln Ala Ser Ala Leu Phe Thr Leu Ala Arg Gly 2345 2350 2355	7170
ttc ccc ttc gtc gat gtt gga gtg tcg gct ctc ctg cta gca gcc gga Phe Pro Phe Val Asp Val Gly Val Ser Ala Leu Leu Leu Ala Ala Gly 2360 2365 2370	7218
tgc tgg gga caa gtc acc ctc acc gtt acg gta aca gcg gca aca ctc Cys Trp Gly Gln Val Thr Leu Thr Val Thr Val Thr Ala Ala Thr Leu 2375 2380 2385 2390	7266
ctt ttt tgc cac tat gcc tac atg gtt ccc ggt tgg caa gct gag gca Leu Phe Cys His Tyr Ala Tyr Met Val Pro Gly Trp Gln Ala Glu Ala 2395 2400 2405	7314
atg cgc tca gcc cag cgg cgg aca gcg gcc gga atc atg aaa aac gct Met Arg Ser Ala Gln Arg Arg Thr Ala Ala Gly Ile Met Lys Asn Ala 2410 2415 2420	7362
gta gtg gat ggc atc gtg gcc acg gac gtc cca gaa tta gag cgc acc Val Val Asp Gly Ile Val Ala Thr Asp Val Pro Glu Leu Glu Arg Thr 2425 2430 2435	7410
aca ccc atc atg cag aag aaa gtt gga cag atc atg ctg atc ttg gtg Thr Pro Ile Met Gln Lys Lys Val Gly Gln Ile Met Leu Ile Leu Val 2440 2445 2450	7458
tct cta gct gca gta gta gtg aac ccg tct gtg aag aca gta cga gaa Ser Leu Ala Ala Val Val Asn Pro Ser Val Lys Thr Val Arg Glu 2455 2460 2465 2470	7506

gcc gga att ttg atc acg gcc gca gcg gtg acg ctt tgg gag aat gga Ala Gly Ile Leu Ile Thr Ala Ala Ala Val Thr Leu Trp Glu Asn Gly 2475 2480 2485	7554
gca agc tct gtt tgg aac gca aca act gcc atc gga ctc tgc cac atc Ala Ser Ser Val Trp Asn Ala Thr Thr Ala Ile Gly Leu Cys His Ile 2490 2495 2500	7602
atg cgt ggg ggt tgg ttg tca tgt cta tcc ata 'aca tgg aca ctc ata Met Arg Gly Gly Trp Leu Ser Cys Leu Ser Ile Thr Trp Thr Leu Ile 2505 2510 2515	7650
aag aac atg gaa aaa cca gga cta aaa aga ggt ggg gca aaa gga cgc Lys Asn Met Glu Lys Pro Gly Leu Lys Arg Gly Gly Ala Lys Gly Arg 2520 2525 2530	7698
acc ttg gga gag gtt tgg aaa gaa aga ctc aac cag atg aca aaa gaa Thr Leu Gly Glu Val Trp Lys Glu Arg Leu Asn Gln Met Thr Lys Glu 2535 2540 2545 2550	7746
gag ttc act agg tac cgc aaa gag gcc atc atc gaa gtc gat cgc tca Glu Phe Thr Arg Tyr Arg Lys Glu Ala Ile Ile Glu Val Asp Arg Ser 2555 2560 2565	7794
gcg gca aaa cac gcc agg aaa gaa ggc aat gtc act gga ggg cat tca Ala Ala Lys His Ala Arg Lys Glu Gly Asn Val Thr Gly Gly His Ser 2570 2575 2580	7842
gtc tct agg ggc aca gca aaa ctg aga tgg ctg gtc gaa cgg agg ttt Val Ser Arg Gly Thr Ala Lys Leu Arg Trp Leu Val Glu Arg Arg Phe 2585 2590 2595	7890
ctc gaa ccg gtc gga aaa gtg att gac ctt gga tgt gga aga ggc ggt Leu Glu Pro Val Gly Lys Val Ile Asp Leu Gly Cys Gly Arg Gly Gly 2600 2605 2610	7938
tgg tgt tac tat atg gca acc caa aaa aga gtc caa gaa gtc aga ggg Trp Cys Tyr Tyr Met Ala Thr Gln Lys Arg Val Gln Glu Val Arg Gly 2615 2620 2625 2630	7986
tac aca aag ggc ggt ccc gga cat gaa gag ccc caa cta gtg caa agt Tyr Thr Lys Gly Gly Pro Gly His Glu Glu Pro Gln Leu Val Gln Ser 2635 2640 2645	8034
tat gga tgg aac att gtc acc atg aag agt gga gtg gat gtg ttc tac Tyr Gly Trp Asn Ile Val Thr Met Lys Ser Gly Val Asp Val Phe Tyr 2650 2655 2660	8082
aga cct tct gag tgt tgt gac acc ctc ctt tgt gac atc gga gag tcc Arg Pro Ser Glu Cys Cys Asp Thr Leu Leu Cys Asp Ile Gly Glu Ser 2665 2670 2675	8130
tcg tca agt gct gag gtt gaa gag cat agg acg att cgt gtc ctt gaa Ser Ser Ser Ala Glu Val Glu Glu His Arg Thr Ile Arg Val Leu Glu 2680 2685 2690	8178
atg gtt gag gac tgg ctg cac cga ggg cca agg gaa ttt tgc gtg aag Met Val Glu Asp Trp Leu His Arg Gly Pro Arg Glu Phe Cys Val Lys 2695 2700 2705 2710	8226

gtg ctc tgc ccc tac atg ccg aaa gtc ata gag aag atg gag ctg ctc Val Leu Cys Pro Tyr Met Pro Lys Val Ile Glu Lys Met Glu Leu Leu 2715 2720 2725	8274
caa cgc cgg tat ggg ggg gga ctg gtc aga aac cca ctc tca cgg aat Gln Arg Arg Tyr Gly Gly Gly Leu Val Arg Asn Pro Leu Ser Arg Asn 2730 2735 2740	8322
tcc acg cac gag atg tat tgg gtg agt cga gct tca ggc aat gtg gta Ser Thr His Glu Met Tyr Trp Val Ser Arg Ala Ser Gly Asn Val Val 2745 2750 2755	8370
cat tca gtg aat atg acc agc cag gtg ctc cta gga aga atg gaa aaa His Ser Val Asn Met Thr Ser Gln Val Leu Leu Gly Arg Met Glu Lys 2760 2765 2770	8418
agg acc tgg aag gga ccc caa tac gag gaa gat gta aac ttg gga agc Arg Thr Trp Lys Gly Pro Gln Tyr Glu Glu Asp Val Asn Leu Gly Ser 2775 2780 2785 2790	8466
gga acc agg gcg gtg gga aaa ccc ctg ctc aac tca gac acc agt aaa Gly Thr Arg Ala Val Gly Lys Pro Leu Leu Asn Ser Asp Thr Ser Lys 2795 2800 2805	8514
atc aac aac agg att gaa cga ctc agg cgt gag tac agt tcg acg tgg Ile Asn Asn Arg Ile Glu Arg Leu Arg Arg Glu Tyr Ser Ser Thr Trp 2810 2815 2820	8562
cac cac gat gag aac cac cca tat aga acc tgg aac tat cac ggc agt His His Asp Glu Asn His Pro Tyr Arg Thr Trp Asn Tyr His Gly Ser 2825 2830 2835	8610
tat gat gtg aag ccc aca ggc tcc gcc agt tcg ctg gtc aat gga gtg Tyr Asp Val Lys Pro Thr Gly Ser Ala Ser Ser Leu Val Asn Gly Val 2840 2845 2850	8658
gtc agg ctc ctc tca aaa cca tgg gac acc atc acg aat gtt acc acc Val Arg Leu Leu Ser Lys Pro Trp Asp Thr Ile Thr Asn Val Thr Thr 2855 2860 2865 2870	8706
atg gcc atg act gac act act ccc ttc ggg cag cag cga gtg ttc aaa Met Ala Met Thr Asp Thr Thr Pro Phe Gly Gln Gln Arg Val Phe Lys 2875 2880 2885	8754
gag aag gtg gac acg aaa gct cct gaa ccg cca gaa gga gcg aag tac Glu Lys Val Asp Thr Lys Ala Pro Glu Pro Pro Glu Gly Ala Lys Tyr 2890 2895 2900	8802
gtg ctc aac gag acc acc aac tgg ttg tgg gcg ttt ttg gcc aga gaa Val Leu Asn Glu Thr Thr Asn Trp Leu Trp Ala Phe Leu Ala Arg Glu 2905 2910 2915	8850
aaa cgt ccc aga atg tgc tct cga gag gaa ttc ata aga aag gtc aac Lys Arg Pro Arg Met Cys Ser Arg Glu Glu Phe Ile Arg Lys Val Asn 2920 2925 2930	8898
agc aat gca gct ttg ggt gcc atg ttt gaa gag cag aat caa tgg agg Ser Asn Ala Ala Leu Gly Ala Met Phe Glu Glu Gln Asn Gln Trp Arg 2935 2940 2945 2950	8946

agc gcc aga gaa gca gtt gaa gat cca aaa ttt tgg gag atg gtg gat	8994
Ser Ala Arg Glu Ala Val Glu Asp Pro Lys Phe Trp Glu Met Val Asp	
2955 2960 2965	
gag gag cgc gag gca cat ctg cgg ggg gaa tgt cac act tgc att tac	9042
Glu Glu Arg Glu Ala His Leu Arg Gly Glu Cys His Thr Cys Ile Tyr	
2970 2975 2980	
aac atg atg gga aag aga gag aaa aaa ccc gga gag ttc gga aag gcc	9090
Asn Met Met Gly Lys Arg Glu Lys Lys Pro Gly Glu Phe Gly Lys Ala	
2985 2990 2995	
aag gga agc aga gcc att tgg ttc atg tgg ctc gga gct cgc ttt ctg	9138
Lys Gly Ser Arg Ala Ile Trp Phe Met Trp Leu Gly Ala Arg Phe Leu	
3000 3005 3010	
gag ttc gag gct ctg ggt ttt ctc aat gaa gac cac tgg ctt gga aga	9186
Glu Phe Glu Ala Leu Gly Phe Leu Asn Glu Asp His Trp Leu Gly Arg	
3015 3020 3025 3030	
aag aac tca gga gga ggt gtc gag ggc ttg ggc ctc caa aaa ctg ggt	9234
Lys Asn Ser Gly Gly Gly Val Glu Gly Leu Gly Leu Gln Lys Leu Gly	
3035 3040 3045	
tac atc ctg cgt gaa gtt ggc acc cgg cct ggg ggc aag atc tat gct	9282
Tyr Ile Leu Arg Glu Val Gly Thr Arg Pro Gly Gly Lys Ile Tyr Ala	
3050 3055 3060	
gat gac aca gct ggc tgg gac acc cgc atc acg aga gct gac ttg gaa	9330
Asp Asp Thr Ala Gly Trp Asp Thr Arg Ile Thr Arg Ala Asp Leu Glu	
3065 3070 3075	
aat gaa gct aag gtg ctt gag ctg ctt gat ggg gaa cat cgg cgt ctt	9378
Asn Glu Ala Lys Val Leu Glu Leu Leu Asp Gly Glu His Arg Arg Leu	
3080 3085 3090	
gcc agg gcc atc att gag ctc acc tat cgt cac aaa gtt gtg aaa gtg	9426
Ala Arg Ala Ile Ile Glu Leu Thr Tyr Arg His Lys Val Val Lys Val	
3095 3100 3105 3110	
atg cgc ccg gct gct gat gga aga acc gtc atg gat gtt atc tcc aga	9474
Met Arg Pro Ala Ala Asp Gly Arg Thr Val Met Asp Val Ile Ser Arg	
3115 3120 3125	
gaa gat cag agg ggg agt gga caa gtt gtc acc tac gcc cta aac act	9522
Glu Asp Gln Arg Gly Ser Gly Gln Val Val Thr Tyr Ala Leu Asn Thr	
3130 3135 3140	
ttc acc aac ctg gcc gtc cag ctg gtg agg atg atg gaa ggg gaa gga	9570
Phe Thr Asn Leu Ala Val Gln Leu Val Arg Met Met Glu Gly Glu Gly	
3145 3150 3155	
gtg att ggc cca gat gat gtg gag aaa ctc aca aaa ggg aaa gga ccc	9618
Val Ile Gly Pro Asp Asp Val Glu Lys Leu Thr Lys Gly Lys Gly Pro	
3160 3165 3170	
aaa gtc agg acc tgg ctg ttt gag aat ggg gaa gaa aga ctc agc cgc	9666
Lys Val Arg Thr Trp Leu Phe Glu Asn Gly Glu Glu Arg Leu Ser Arg	
3175 3180 3185 3190	

atg gct gtc agt gga gat gac tgt gtg gta aag ccc ctg gac gat cgc Met Ala Val Ser Gly Asp Asp Cys Val Val Lys Pro Leu Asp Asp Arg 3195 3200 3205	9714
ttt gcc acc tcg ctc cac ttc ctc aat gct atg tca aag gtt cgc aaa Phe Ala Thr Ser Leu His Phe Leu Asn Ala Met Ser Lys Val Arg Lys 3210 3215 3220	9762
gac atc caa gag tgg aaa ccg tca act gga tgg tat gat tgg cag cag Asp Ile Gln Glu Trp Lys Pro Ser Thr Gly Trp Tyr Asp Trp Gln Gln 3225 3230 3235	9810
gtt cca ttt tgc tca aac cat ttc act gaa ttg atc atg aaa gat gga Val Pro Phe Cys Ser Asn His Phe Thr Glu Leu Ile Met Lys Asp Gly 3240 3245 3250	9858
aga aca ctg gtg gtt cca tgc cga gga cag gat gaa ttg gta ggc aga Arg Thr Leu Val Val Pro Cys Arg Gly Gln Asp Glu Leu Val Gly Arg 3255 3260 3265 3270	9906
gct cgc ata tct cca ggg gcc gga tgg aac gtc cgc gac act gct tgt Ala Arg Ile Ser Pro Gly Ala Gly Trp Asn Val Arg Asp Thr Ala Cys 3275 3280 3285	9954
ctg gct aag tct tat gcc cag atg tgg ctg ctt ctg tac ttc cac aga Leu Ala Lys Ser Tyr Ala Gln Met Trp Leu Leu Leu Tyr Phe His Arg 3290 3295 3300	10002
aga gac ctg cgg ctc atg gcc aac gcc att tgc tcc gct gtc cct gtg Arg Asp Leu Arg Leu Met Ala Asn Ala Ile Cys Ser Ala Val Pro Val 3305 3310 3315	10050
aat tgg gtc cct acc gga aga acc acg tgg tcc atc cat gca gga gga Asn Trp Val Pro Thr Gly Arg Thr Thr Trp Ser Ile His Ala Gly Gly 3320 3325 3330	10098
gag tgg atg aca aca gag gac atg ttg gag gtc tgg aac cgt gtt tgg Glu Trp Met Thr Thr Glu Asp Met Leu Glu Val Trp Asn Arg Val Trp 3335 3340 3345 3350	10146
ata gag gag aat gaa tgg atg gaa gac aaa acc cca gtg gag aaa tgg Ile Glu Glu Asn Glu Trp Met Glu Asp Lys Thr Pro Val Glu Lys Trp 3355 3360 3365	10194
agt gac gtc cca tat tca gga aaa cga gag gac atc tgg tgt ggc agc Ser Asp Val Pro Tyr Ser Gly Lys Arg Glu Asp Ile Trp Cys Gly Ser 3370 3375 3380	10242
ctg att ggc aca aga gcc cga gcc acg tgg gca gaa aac atc cag gtg Leu Ile Gly Thr Arg Ala Arg Ala Thr Trp Ala Glu Asn Ile Gln Val 3385 3390 3395	10290
gct atc aac caa gtc aga gca atc atc gga gat gag aag tat gtg gat Ala Ile Asn Gln Val Arg Ala Ile Ile Gly Asp Glu Lys Tyr Val Asp 3400 3405 3410	10338
tac atg agt tca cta aag aga tat gaa gac aca act ttg gtt gag gac Tyr Met Ser Ser Leu Lys Arg Tyr Glu Asp Thr Thr Leu Val Glu Asp 3415 3420 3425 3430	10386

```

aca gta ctg tagatattta atcaattgta aatagacaat ataagtatgc      10435
Thr Val Leu

ataaaaagtgt agttttatag tagtatttag tgggtgtagt gtaaatagtt aagaaaattt 10495
tgaggagaaa gtcaggccgg gaagttcccg ccaccggaag ttgagtagac ggtgctgcct 10555
gcgactcaac cccaggagga ctgggtgaac aaagccgcga agtgatccat gtaagccctc 10615
agaaccgtct cggaaggagg accccacatg ttgtaacttc aaagcccaat gtcagaccac 10675
gctacggcgt gctactctgc ggagagtgcg gtctgcgata gtgccccagg aggactgggt 10735
taacaaaggc aaaccaacgc cccacgcggc cctagccccg gtaatggtgt taaccagggc 10795
gaaaggacta gaggttagag gagacccgcg ggtttaaaagt gcacggccca gcctgactga 10855
agctgtaggt caggggaagg actagagggt agtggagacc ccgtgccaca aaacaccaca 10915
acaaaacagc atattgacac ctgggataga ctaggagatc ttctgctctg cacaaccagc 10975
cacacggcac agtgcgccga caatggtggc tgggtggtgcg agaacacagg atct      11029

```

< 210 > 2

< 211 > 3433

< 212 > PRT

5 < 213 > Flavivirus sp.

< 400 > 2

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

Ala Met Asp Val Gly Tyr Met Cys Asp Asp Thr Ile Thr Tyr Glu Cys
 165 170 175
 Pro Val Leu Ser Ala Gly Asn Asp Pro Glu Asp Ile Asp Cys Trp Cys
 180 185 190
 Thr Lys Ser Ala Val Tyr Val Arg Tyr Gly Arg Cys Thr Lys Thr Arg
 195 200 205
 His Ser Arg Arg Ser Arg Arg Ser Leu Thr Val Gln Thr His Gly Glu
 210 215 220
 Ser Thr Leu Ala Asn Lys Lys Gly Ala Trp Met Asp Ser Thr Lys Ala
 225 230 235 240
 Thr Arg Tyr Leu Val Lys Thr Glu Ser Trp Ile Leu Arg Asn Pro Gly
 245 250 255
 Tyr Ala Leu Val Ala Ala Val Ile Gly Trp Met Leu Gly Ser Asn Thr
 260 265 270
 Met Gln Arg Val Val Phe Val Val Leu Leu Leu Val Ala Pro Ala
 275 280 285
 Tyr Ser Phe Asn Cys Leu Gly Met Ser Asn Arg Asp Phe Leu Glu Gly
 290 295 300
 Val Ser Gly Ala Thr Trp Val Asp Leu Val Leu Glu Gly Asp Ser Cys
 305 310 315 320
 Val Thr Ile Met Ser Lys Asp Lys Pro Thr Ile Asp Val Lys Met Met
 325 330 335
 Asn Met Glu Ala Ala Asn Leu Ala Glu Val Arg Ser Tyr Cys Tyr Leu
 340 345 350
 Ala Thr Val Ser Asp Leu Ser Thr Lys Ala Ala Cys Pro Thr Met Gly
 355 360 365
 Glu Ala His Asn Asp Lys Arg Ala Asp Pro Ala Phe Val Cys Arg Gln
 370 375 380
 Gly Val Val Asp Arg Gly Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys
 385 390 395 400
 Gly Ser Ile Asp Thr Cys Ala Lys Phe Ala Cys Ser Thr Lys Ala Ile
 405 410 415
 Gly Arg Thr Ile Leu Lys Glu Asn Ile Lys Tyr Glu Val Ala Ile Phe
 420 425 430
 Val His Gly Pro Thr Thr Val Glu Ser His Gly Asn Tyr Ser Thr Gln
 435 440 445
 Val Gly Ala Thr Gln Ala Gly Arg Phe Ser Ile Thr Pro Ala Ala Pro
 450 455 460
 Ser Tyr Thr Leu Lys Leu Gly Glu Tyr Gly Glu Val Thr Val Asp Cys
 465 470 475 480

Glu Pro Arg Ser Gly Ile Asp Thr Asn Ala Tyr Tyr Val Met Thr Val
 485 490 495
 Gly Thr Lys Thr Phe Leu Val His Arg Glu Trp Phe Met Asp Leu Asn
 500 505 510
 Leu Pro Trp Ser Ser Ala Gly Ser Thr Val Trp Arg Asn Arg Glu Thr
 515 520 525
 Leu Met Glu Phe Glu Glu Pro His Ala Thr Lys Gln Ser Val Ile Ala
 530 535 540
 Leu Gly Ser Gln Glu Gly Ala Leu His Gln Ala Leu Ala Gly Ala Ile
 545 550 555 560
 Pro Val Glu Phe Ser Ser Asn Thr Val Lys Leu Thr Ser Gly His Leu
 565 570 575
 Lys Cys Arg Val Lys Met Glu Lys Leu Gln Leu Lys Gly Thr Thr Tyr
 580 585 590
 Gly Val Cys Ser Lys Ala Phe Lys Phe Leu Gly Thr Pro Ala Asp Thr
 595 600 605
 Gly His Gly Thr Val Val Leu Glu Leu Gln Tyr Thr Gly Thr Asp Gly
 610 615 620
 Pro Cys Lys Val Pro Ile Ser Ser Val Ala Ser Leu Asn Asp Leu Thr
 625 630 635 640
 Pro Val Gly Arg Leu Val Thr Val Asn Pro Phe Val Ser Val Ala Thr
 645 650 655
 Ala Asn Ala Lys Val Leu Ile Glu Leu Glu Pro Pro Phe Gly Asp Ser
 660 665 670
 Tyr Ile Val Val Gly Arg Gly Glu Gln Gln Ile Asn His His Trp His
 675 680 685
 Lys Ser Gly Ser Ser Ile Gly Lys Ala Phe Thr Thr Thr Leu Lys Gly
 690 695 700
 Ala Gln Arg Leu Ala Ala Leu Gly Asp Thr Ala Trp Asp Phe Gly Ser
 705 710 715 720
 Val Gly Gly Val Phe Thr Ser Val Gly Lys Ala Val His Gln Val Phe
 725 730 735
 Gly Gly Ala Phe Arg Ser Leu Phe Gly Gly Met Ser Trp Ile Thr Gln
 740 745 750
 Gly Leu Leu Gly Ala Leu Leu Leu Trp Met Gly Ile Asn Ala Arg Asp
 755 760 765
 Arg Ser Ile Ala Leu Thr Phe Leu Ala Val Gly Gly Val Leu Leu Phe
 770 775 780
 Leu Ser Val Asn Val His Ala Asp Thr Gly Cys Ala Ile Asp Ile Ser
 785 790 795 800

Arg Gln Glu Leu Arg Cys Gly Asn Gly Val Phe Ile His Asn Asp Val
 805 810 815
 Glu Ala Trp Met Asp Arg Tyr Lys Tyr Tyr Pro Glu Thr Pro Gln Gly
 820 825 830
 Leu Ala Lys Ile Ile Gln Lys Ala His Lys Glu Gly Val Cys Gly Leu
 835 840 845
 Arg Ser Val Ser Arg Leu Glu His Gln Met Trp Glu Ala Val Lys Asp
 850 855 860
 Glu Leu Asn Thr Leu Leu Lys Glu Asn Gly Val Asp Leu Ser Val Val
 865 870 875 880
 Val Glu Lys Gln Glu Gly Met Tyr Lys Ser Ala Pro Lys Arg Leu Thr
 885 890 895
 Ala Thr Thr Glu Lys Leu Glu Ile Gly Trp Lys Ala Trp Gly Lys Ser
 900 905 910
 Ile Leu Phe Ala Pro Glu Leu Ala Asn Asn Thr Phe Val Val Asp Gly
 915 920 925
 Pro Glu Thr Lys Glu Cys Pro Thr Gln Asn Arg Ala Trp Asn Ser Leu
 930 935 940
 Glu Val Glu Asp Phe Gly Phe Gly Leu Thr Ser Thr Arg Met Phe Leu
 945 950 955 960
 Lys Val Arg Glu Ser Asn Thr Thr Glu Cys Asp Ser Lys Ile Ile Gly
 965 970 975
 Thr Ala Val Lys Asn Asn Leu Ala Ile His Ser Asp Leu Ser Tyr Trp
 980 985 990
 Ile Glu Ser Arg Leu Asn Asp Thr Trp Lys Leu Glu Arg Ala Val Leu
 995 1000 1005
 Gly Glu Val Lys Ser Cys Thr Trp Pro Glu Thr His Thr Leu Trp Gly
 1010 1015 1020
 Asp Gly Ile Leu Glu Ser Asp Leu Ile Ile Pro Val Thr Leu Ala Gly
 1025 1030 1035 1040
 Pro Arg Ser Asn His Asn Arg Arg Pro Gly Tyr Lys Thr Gln Asn Gln
 1045 1050 1055
 Gly Pro Trp Asp Glu Gly Arg Val Glu Ile Asp Phe Asp Tyr Cys Pro
 1060 1065 1070
 Gly Thr Thr Val Thr Leu Ser Glu Ser Cys Gly His Arg Gly Pro Ala
 1075 1080 1085
 Thr Arg Thr Thr Thr Glu Ser Gly Lys Leu Ile Thr Asp Trp Cys Cys
 1090 1095 1100
 Arg Ser Cys Thr Leu Pro Pro Leu Arg Tyr Gln Thr Asp Ser Gly Cys
 1105 1110 1115 1120

Trp Tyr Gly Met Glu Ile Arg Pro Gln Arg His Asp Glu Lys Thr Leu
 1125 1130 1135
 Val Gln Ser Gln Val Asn Ala Tyr Asn Ala Asp Met Ile Asp Pro Phe
 1140 1145 1150
 Gln Leu Gly Leu Leu Val Val Phe Leu Ala Thr Gln Glu Val Leu Arg
 1155 1160 1165
 Lys Arg Trp Thr Ala Lys Ile Ser Met Pro Ala Ile Leu Ile Ala Leu
 1170 1175 1180
 Leu Val Leu Val Phe Gly Gly Ile Thr Tyr Thr Asp Val Leu Arg Tyr
 1185 1190 1195 1200
 Val Ile Leu Val Gly Ala Ala Phe Ala Glu Ser Asn Ser Gly Gly Asp
 1205 1210 1215
 Val Val His Leu Ala Leu Met Ala Thr Phe Lys Ile Gln Pro Val Phe
 1220 1225 1230
 Met Val Ala Ser Phe Leu Lys Ala Arg Trp Thr Asn Gln Glu Asn Ile
 1235 1240 1245
 Leu Leu Met Leu Ala Ala Val Phe Phe Gln Met Ala Tyr His Asp Ala
 1250 1255 1260
 Arg Gln Ile Leu Leu Trp Glu Ile Pro Asp Val Leu Asn Ser Leu Ala
 1265 1270 1275 1280
 Val Ala Trp Met Ile Leu Arg Ala Ile Thr Phe Thr Thr Thr Ser Asn
 1285 1290 1295
 Val Val Val Pro Leu Leu Ala Leu Leu Thr Pro Arg Leu Arg Cys Leu
 1300 1305 1310
 Asn Leu Asp Val Tyr Arg Ile Leu Leu Leu Met Val Gly Ile Gly Ser
 1315 1320 1325
 Leu Ile Arg Glu Lys Arg Ser Ala Ala Ala Lys Lys Lys Gly Ala Ser
 1330 1335 1340
 Leu Leu Cys Leu Ala Leu Ala Ser Thr Gly Leu Phe Asn Pro Met Ile
 1345 1350 1355 1360
 Leu Ala Ala Gly Leu Ile Ala Cys Asp Pro Asn Arg Lys Arg Gly Trp
 1365 1370 1375
 Pro Ala Thr Glu Val Met Thr Ala Val Gly Leu Met Phe Ala Ile Val
 1380 1385 1390
 Gly Gly Leu Ala Glu Leu Asp Ile Asp Ser Met Ala Ile Pro Met Thr
 1395 1400 1405
 Ile Ala Gly Leu Met Phe Ala Ala Phe Val Ile Ser Gly Lys Ser Thr
 1410 1415 1420
 Asp Met Trp Ile Glu Arg Thr Ala Asp Ile Ser Trp Glu Ser Asp Ala
 1425 1430 1435 1440

Glu Ile Thr Gly Ser Ser Glu Arg Val Asp Val Arg Leu Asp Asp Gly
 1445 1450 1455
 Glu Asn Phe Gln Leu Met Asn Asp Pro Gly Ala Pro Trp Lys Ile Trp
 1460 1465 1470
 Met Leu Arg Met Val Cys Leu Ala Ile Ser Ala Tyr Thr Pro Trp Ala
 1475 1480 1485
 Ile Leu Pro Ser Val Val Gly Phe Trp Ile Thr Leu Gln Tyr Thr Lys
 1490 1495 1500
 Arg Gly Gly Val Leu Trp Asp Thr Pro Ser Pro Lys Glu Tyr Lys Lys
 505 1510 1515 1520
 Gly Asp Thr Thr Thr Gly Val Tyr Arg Ile Met Thr Arg Gly Leu Leu
 1525 1530 1535
 Gly Ser Tyr Gln Ala Gly Ala Gly Val Met Val Glu Gly Val Phe His
 1540 1545 1550
 Thr Leu Trp His Thr Thr Lys Gly Ala Ala Leu Met Ser Gly Glu Gly
 1555 1560 1565
 Arg Leu Asp Pro Tyr Trp Gly Ser Val Lys Glu Asp Arg Leu Cys Tyr
 1570 1575 1580
 Gly Gly Pro Trp Lys Leu Gln His Lys Trp Asn Gly Gln Asp Glu Val
 585 1590 1595 1600
 Gln Met Ile Val Val Glu Pro Gly Lys Asn Val Lys Asn Val Gln Thr
 1605 1610 1615
 Lys Pro Gly Val Phe Lys Thr Pro Glu Gly Glu Ile Gly Ala Val Thr
 1620 1625 1630
 Leu Asp Phe Pro Thr Gly Thr Ser Gly Ser Pro Ile Val Asp Lys Asn
 1635 1640 1645
 Gly Asp Val Ile Gly Leu Tyr Gly Asn Gly Val Ile Met Pro Asn Gly
 1650 1655 1660
 Ser Tyr Ile Ser Ala Ile Val Gln Gly Glu Arg Met Asp Glu Pro Ile
 665 1670 1675 1680
 Pro Ala Gly Phe Glu Pro Glu Met Leu Arg Lys Lys Gln Ile Thr Val
 1685 1690 1695
 Leu Asp Leu His Pro Gly Ala Gly Lys Thr Arg Arg Ile Leu Pro Gln
 1700 1705 1710
 Ile Ile Lys Glu Ala Ile Asn Arg Arg Leu Arg Thr Ala Val Leu Ala
 1715 1720 1725
 Pro Thr Arg Val Val Ala Ala Glu Met Ala Glu Ala Leu Arg Gly Leu
 1730 1735 1740
 Pro Ile Arg Tyr Gln Thr Ser Ala Val Pro Arg Glu His Asn Gly Asn
 745 1750 1755 1760

Glu Ile Val Asp Val Met Cys His Ala Thr Leu Thr His Arg Leu Met
 1765 1770 1775
 Ser Pro His Arg Val Pro Asn Tyr Asn Leu Phe Val Met Asp Glu Ala
 1780 1785 1790
 His Phe Thr Asp Pro Ala Ser Ile Ala Ala Arg Gly Tyr Ile Ser Thr
 1795 1800 1805
 Lys Val Glu Leu Gly Glu Ala Ala Ala Ile Phe Met Thr Ala Thr Pro
 1810 1815 1820
 Pro Gly Thr Ser Asp Pro Phe Pro Glu Ser Asn Ser Pro Ile Ser Asp
 825 1830 1835 1840
 Leu Gln Thr Glu Ile Pro Asp Arg Ala Trp Asn Ser Gly Tyr Glu Trp
 1845 1850 1855
 Ile Thr Glu Tyr Thr Gly Lys Thr Val Trp Phe Val Pro Ser Val Lys
 1860 1865 1870
 Met Gly Asn Glu Ile Ala Leu Cys Leu Gln Arg Ala Gly Lys Lys Val
 1875 1880 1885
 Val Gln Leu Asn Arg Lys Ser Tyr Glu Thr Glu Tyr Pro Lys Cys Lys
 1890 1895 1900
 Asn Asp Asp Trp Asp Phe Val Ile Thr Thr Asp Ile Ser Glu Met Gly
 905 1910 1915 1920
 Ala Asn Phe Lys Ala Ser Arg Val Ile Asp Ser Arg Lys Ser Val Lys
 1925 1930 1935
 Pro Thr Ile Ile Thr Glu Gly Glu Gly Arg Val Ile Leu Gly Glu Pro
 1940 1945 1950
 Ser Ala Val Thr Ala Ala Ser Ala Ala Gln Arg Arg Gly Arg Ile Gly
 1955 1960 1965
 Arg Asn Pro Ser Gln Val Gly Asp Glu Tyr Cys Tyr Gly Gly His Thr
 1970 1975 1980
 Asn Glu Asp Asp Ser Asn Phe Ala His Trp Thr Glu Ala Arg Ile Met
 985 1990 1995 2000
 Pro Asp Asn Ile Asn Met Pro Asn Gly Leu Ile Ala Gln Phe Tyr Gln
 2005 2010 2015
 Pro Glu Arg Glu Lys Val Tyr Thr Met Glu Gly Glu Tyr Arg Leu Arg
 2020 2025 2030
 Gly Glu Glu Arg Lys Asn Phe Leu Glu Leu Leu Arg Thr Ala Asp Leu
 2035 2040 2045
 Pro Val Trp Leu Ala Tyr Lys Val Ala Ala Ala Gly Val Ser Tyr His
 2050 2055 2060
 Asp Arg Arg Trp Cys Phe Asp Gly Pro Arg Thr Asn Thr Ile Leu Glu
 065 2070 2075 2080

Asp Asn Asn Glu Val Glu Val Ile Thr Lys Leu Gly Glu Arg Lys Ile
 2085 2090 2095
 Leu Arg Pro Arg Trp Ile Asp Ala Arg Val Tyr Ser Asp His Gln Ala
 2100 2105 2110
 Leu Lys Ala Phe Lys Asp Phe Ala Ser Gly Lys Arg Ser Gln Ile Gly
 2115 2120 2125
 Leu Ile Glu Val Leu Gly Lys Met Pro Glu His Phe Met Gly Lys Thr
 2130 2135 2140
 Trp Glu Ala Leu Asp Thr Met Tyr Val Val Ala Thr Ala Glu Lys Gly
 145 2150 2155 2160
 Gly Arg Ala His Arg Met Ala Leu Glu Glu Leu Pro Asp Ala Leu Gln
 2165 2170 2175
 Thr Ile Ala Leu Ile Ala Leu Leu Ser Val Met Thr Met Gly Val Phe
 2180 2185 2190
 Phe Leu Leu Met Gln Arg Lys Gly Ile Gly Lys Ile Gly Leu Gly Gly
 2195 2200 2205
 Ala Val Leu Gly Val Ala Thr Phe Phe Cys Trp Met Ala Glu Val Pro
 2210 2215 2220
 Gly Thr Lys Ile Ala Gly Met Leu Leu Leu Ser Leu Leu Leu Met Ile
 225 2230 2235 2240
 Val Leu Ile Pro Glu Pro Glu Lys Gln Arg Ser Gln Thr Asp Asn Gln
 2245 2250 2255
 Leu Ala Val Phe Leu Ile Cys Val Met Thr Leu Val Ser Ala Val Ala
 2260 2265 2270
 Ala Asn Glu Met Gly Trp Leu Asp Lys Thr Lys Ser Asp Ile Ser Ser
 2275 2280 2285
 Leu Phe Gly Gln Arg Ile Glu Val Lys Glu Asn Phe Ser Met Gly Glu
 2290 2295 2300
 Phe Leu Leu Asp Leu Arg Pro Ala Thr Ala Trp Ser Leu Tyr Ala Val
 305 2310 2315 2320
 Thr Thr Ala Val Leu Thr Pro Leu Leu Lys His Leu Ile Thr Ser Asp
 2325 2330 2335
 Tyr Ile Asn Thr Ser Leu Thr Ser Ile Asn Val Gln Ala Ser Ala Leu
 2340 2345 2350
 Phe Thr Leu Ala Arg Gly Phe Pro Phe Val Asp Val Gly Val Ser Ala
 2355 2360 2365
 Leu Leu Leu Ala Ala Gly Cys Trp Gly Gln Val Thr Leu Thr Val Thr
 2370 2375 2380
 Val Thr Ala Ala Thr Leu Leu Phe Cys His Tyr Ala Tyr Met Val Pro
 385 2390 2395 2400

Gly Trp Gln Ala Glu Ala Met Arg Ser Ala Gln Arg Arg Thr Ala Ala
 2405 2410 2415
 Gly Ile Met Lys Asn Ala Val Val Asp Gly Ile Val Ala Thr Asp Val
 2420 2425 2430
 Pro Glu Leu Glu Arg Thr Thr Pro Ile Met Gln Lys Lys Val Gly Gln
 2435 2440 2445
 Ile Met Leu Ile Leu Val Ser Leu Ala Ala Val Val Val Asn Pro Ser
 2450 2455 2460
 Val Lys Thr Val Arg Glu Ala Gly Ile Leu Ile Thr Ala Ala Ala Val
 465 2470 2475 2480
 Thr Leu Trp Glu Asn Gly Ala Ser Ser Val Trp Asn Ala Thr Thr Ala
 2485 2490 2495
 Ile Gly Leu Cys His Ile Met Arg Gly Gly Trp Leu Ser Cys Leu Ser
 2500 2505 2510
 Ile Thr Trp Thr Leu Ile Lys Asn Met Glu Lys Pro Gly Leu Lys Arg
 2515 2520 2525
 Gly Gly Ala Lys Gly Arg Thr Leu Gly Glu Val Trp Lys Glu Arg Leu
 2530 2535 2540
 Asn Gln Met Thr Lys Glu Glu Phe Thr Arg Tyr Arg Lys Glu Ala Ile
 545 2550 2555 2560
 Ile Glu Val Asp Arg Ser Ala Ala Lys His Ala Arg Lys Glu Gly Asn
 2565 2570 2575
 Val Thr Gly Gly His Ser Val Ser Arg Gly Thr Ala Lys Leu Arg Trp
 2580 2585 2590
 Leu Val Glu Arg Arg Phe Leu Glu Pro Val Gly Lys Val Ile Asp Leu
 2595 2600 2605
 Gly Cys Gly Arg Gly Gly Trp Cys Tyr Tyr Met Ala Thr Gln Lys Arg
 2610 2615 2620
 Val Gln Glu Val Arg Gly Tyr Thr Lys Gly Gly Pro Gly His Glu Glu
 625 2630 2635 2640
 Pro Gln Leu Val Gln Ser Tyr Gly Trp Asn Ile Val Thr Met Lys Ser
 2645 2650 2655
 Gly Val Asp Val Phe Tyr Arg Pro Ser Glu Cys Cys Asp Thr Leu Leu
 2660 2665 2670
 Cys Asp Ile Gly Glu Ser Ser Ser Ser Ala Glu Val Glu Glu His Arg
 2675 2680 2685
 Thr Ile Arg Val Leu Glu Met Val Glu Asp Trp Leu His Arg Gly Pro
 2690 2695 2700
 Arg Glu Phe Cys Val Lys Val Leu Cys Pro Tyr Met Pro Lys Val Ile
 705 2710 2715 2720

Glu Lys Met Glu Leu Leu Gln Arg Arg Tyr Gly Gly Gly Leu Val Arg
 2725 2730 2735
 Asn Pro Leu Ser Arg Asn Ser Thr His Glu Met Tyr Trp Val Ser Arg
 2740 2745 2750
 Ala Ser Gly Asn Val Val His Ser Val Asn Met Thr Ser Gln Val Leu
 2755 2760 2765
 Leu Gly Arg Met Glu Lys Arg Thr Trp Lys Gly Pro Gln Tyr Glu Glu
 2770 2775 2780
 Asp Val Asn Leu Gly Ser Gly Thr Arg Ala Val Gly Lys Pro Leu Leu
 785 2790 2795 2800
 Asn Ser Asp Thr Ser Lys Ile Asn Asn Arg Ile Glu Arg Leu Arg Arg
 2805 2810 2815
 Glu Tyr Ser Ser Thr Trp His His Asp Glu Asn His Pro Tyr Arg Thr
 2820 2825 2830
 Trp Asn Tyr His Gly Ser Tyr Asp Val Lys Pro Thr Gly Ser Ala Ser
 2835 2840 2845
 Ser Leu Val Asn Gly Val Val Arg Leu Leu Ser Lys Pro Trp Asp Thr
 2850 2855 2860
 Ile Thr Asn Val Thr Thr Met Ala Met Thr Asp Thr Thr Pro Phe Gly
 865 2870 2875 2880
 Gln Gln Arg Val Phe Lys Glu Lys Val Asp Thr Lys Ala Pro Glu Pro
 2885 2890 2895
 Pro Glu Gly Ala Lys Tyr Val Leu Asn Glu Thr Thr Asn Trp Leu Trp
 2900 2905 2910
 Ala Phe Leu Ala Arg Glu Lys Arg Pro Arg Met Cys Ser Arg Glu Glu
 2915 2920 2925
 Phe Ile Arg Lys Val Asn Ser Asn Ala Ala Leu Gly Ala Met Phe Glu
 2930 2935 2940
 Glu Gln Asn Gln Trp Arg Ser Ala Arg Glu Ala Val Glu Asp Pro Lys
 945 2950 2955 2960
 Phe Trp Glu Met Val Asp Glu Glu Arg Glu Ala His Leu Arg Gly Glu
 2965 2970 2975
 Cys His Thr Cys Ile Tyr Asn Met Met Gly Lys Arg Glu Lys Lys Pro
 2980 2985 2990
 Gly Glu Phe Gly Lys Ala Lys Gly Ser Arg Ala Ile Trp Phe Met Trp
 2995 3000 3005
 Leu Gly Ala Arg Phe Leu Glu Phe Glu Ala Leu Gly Phe Leu Asn Glu
 3010 3015 3020
 Asp His Trp Leu Gly Arg Lys Asn Ser Gly Gly Gly Val Glu Gly Leu
 025 3030 3035 3040

Gly Leu Gln Lys Leu Gly Tyr Ile Leu Arg Glu Val Gly Thr Arg Pro
 3045 3050 3055
 Gly Gly Lys Ile Tyr Ala Asp Asp Thr Ala Gly Trp Asp Thr Arg Ile
 3060 3065 3070
 Thr Arg Ala Asp Leu Glu Asn Glu Ala Lys Val Leu Glu Leu Leu Asp
 3075 3080 3085
 Gly Glu His Arg Arg Leu Ala Arg Ala Ile Ile Glu Leu Thr Tyr Arg
 3090 3095 3100
 His Lys Val Val Lys Val Met Arg Pro Ala Ala Asp Gly Arg Thr Val
 105 3110 3115 3120
 Met Asp Val Ile Ser Arg Glu Asp Gln Arg Gly Ser Gly Gln Val Val
 3125 3130 3135
 Thr Tyr Ala Leu Asn Thr Phe Thr Asn Leu Ala Val Gln Leu Val Arg
 3140 3145 3150
 Met Met Glu Gly Glu Gly Val Ile Gly Pro Asp Asp Val Glu Lys Leu
 3155 3160 3165
 Thr Lys Gly Lys Gly Pro Lys Val Arg Thr Trp Leu Phe Glu Asn Gly
 3170 3175 3180
 Glu Glu Arg Leu Ser Arg Met Ala Val Ser Gly Asp Asp Cys Val Val
 185 3190 3195 3200
 Lys Pro Leu Asp Asp Arg Phe Ala Thr Ser Leu His Phe Leu Asn Ala
 3205 3210 3215
 Met Ser Lys Val Arg Lys Asp Ile Gln Glu Trp Lys Pro Ser Thr Gly
 3220 3225 3230
 Trp Tyr Asp Trp Gln Gln Val Pro Phe Cys Ser Asn His Phe Thr Glu
 3235 3240 3245
 Leu Ile Met Lys Asp Gly Arg Thr Leu Val Val Pro Cys Arg Gly Gln
 3250 3255 3260
 Asp Glu Leu Val Gly Arg Ala Arg Ile Ser Pro Gly Ala Gly Trp Asn
 265 3270 3275 3280
 Val Arg Asp Thr Ala Cys Leu Ala Lys Ser Tyr Ala Gln Met Trp Leu
 3285 3290 3295
 Leu Leu Tyr Phe His Arg Arg Asp Leu Arg Leu Met Ala Asn Ala Ile
 3300 3305 3310
 Cys Ser Ala Val Pro Val Asn Trp Val Pro Thr Gly Arg Thr Thr Trp
 3315 3320 3325
 Ser Ile His Ala Gly Gly Glu Trp Met Thr Thr Glu Asp Met Leu Glu
 3330 3335 3340
 Val Trp Asn Arg Val Trp Ile Glu Glu Asn Glu Trp Met Glu Asp Lys
 345 3350 3355 3360

Thr Pro Val Glu Lys Trp Ser Asp Val Pro Tyr Ser Gly Lys Arg Glu
 3365 3370 3375
 Asp Ile Trp Cys Gly Ser Leu Ile Gly Thr Arg Ala Arg Ala Thr Trp
 3380 3385 3390
 Ala Glu Asn Ile Gln Val Ala Ile Asn Gln Val Arg Ala Ile Ile Gly
 3395 3400 3405
 Asp Glu Lys Tyr Val Asp Tyr Met Ser Ser Leu Lys Arg Tyr Glu Asp
 3410 3415 3420
 Thr Thr Leu Val Glu Asp Thr Val Leu
 425 3430

< 210 > 3

< 211 > 37

5 < 212 > ADN

< 213 > Secuencia artificial

< 220 >

< 223 > Descripción de la secuencia artificial: CEBADOR OLIGONUCLEOTÍDICO

10

< 400 > 3

tagcacgaag aattcgatgt ctaagaaacc aggaggg 37

15 < 210 > 4

< 211 > 50

< 212 > ADN

< 213 > Secuencia artificial

20 < 220 >

< 223 > Descripción de la secuencia artificial: CEBADOR OLIGONUCLEOTÍDICO

< 400 > 4

25 aagttagccc gggtaatgc tcctacgctg gcgatcaggc caatcaggac 50

< 210 > 5

< 211 > 28

< 212 > ADN

30 < 213 > Secuencia artificial

< 220 >

< 223 > Descripción de la secuencia artificial: CEBADOR

35 < 400 > 5

agtagttcgc ctgtgtgagc tgacaaac 28

< 210 > 6

40 < 211 > 29

< 212 > ADN

< 213 > Secuencia artificial

< 220 >

< 223 > Descripción de la secuencia artificial: cebador

< 400 > 6

5 agatcctgtg ttctgcacc accagccac 29

< 210 > 7

< 211 > 21

< 212 > ADN

10 < 213 > Secuencia artificial

< 220 >

< 223 > Descripción de la secuencia artificial: cebador

15 < 400 > 7

ggatggatgctwggkagc aac 21

< 210 > 8

20 < 211 > 19

< 212 > ADN

< 213 > Secuencia artificial

< 220 >

25 < 223 > Descripción de la secuencia artificial: cebador

< 400 > 8

ccatccaagc ctccacatc 19

30

< 210 > 9

< 211 > 30

< 212 > ADN

< 213 > Secuencia artificial

35

< 220 >

< 223 > Descripción de la secuencia artificial: cebador

< 400 > 9

40

tggatgggat ccaatatgcg tgataggtcc 30

< 210 > 10

< 211 > 39

45 < 212 > ADN

< 213 > Secuencia artificial

< 220 >

< 223 > Descripción de la secuencia artificial: cebador

50

< 400 > 10

aaaaggtca atggtaccag cattttaagc attcacgtt 39

55 < 210 > 11

< 211 > 36

< 212 > ADN

< 213 > Secuencia artificial

< 220 >

< 223 > Descripción de la secuencia artificial: cebador

< 400 > 11

5

tagcacgaagaattcgatgtctaaaaaccaggagg 36

REIVINDICACIONES

1. Cepa aislada del virus West-Nile, **caracterizada porque** su genoma está constituido por la secuencia SEQ ID N° 1.
2. Molécula de ácido nucleico, **caracterizada porque** se selecciona entre el grupo constituido por la secuencia SEQ ID N° 1 y la secuencia complementaria anti-sentido de la secuencia SEQ ID N° 1.
3. Fragmento de la molécula de ácido nucleico de secuencia SEQ ID N° 1, **caracterizado porque** se selecciona entre el grupo constituido por los cebadores que comprenden al menos 15 nucleótidos de dicha secuencia situados entre 10 y 100 nucleótidos más arriba o más abajo de uno de los codones correspondientes a las siguientes posiciones en la secuencia nucleotídica SEQ ID N° 1:
 - codón Asparagina en posiciones 2518-2520 (posición 17 de la secuencia de aminoácidos de la proteína NS1),
 - codón Arginina en posiciones 4018-4020 (posición 164 de la secuencia de aminoácidos de la proteína NS2A),
 - codones Glicina en posiciones 4462-4464 y ácido glutámico en posiciones 4465-4467 (posiciones 82 y 83 de la secuencia de aminoácidos de la proteína NS2B),
 - codones Prolina en posiciones 6097-6099 y ácido glutámico en posiciones 6172-6174 (posiciones 496 y 521 de la secuencia de aminoácidos de la proteína NS3), y
 - codones Serina en posiciones 7840-7842, Asparagina en posiciones 8518-8520 y Alanina en posiciones 8794-8796 (posiciones 54, 280 y 372 de la secuencia de aminoácidos de la proteína NS5).
4. Fragmento de la molécula de ácido nucleico de secuencia SEQ ID N° 1, **caracterizado porque** se selecciona entre el grupo constituido por los fragmentos, preferentemente entre 50 y 200 nucleótidos, amplificados usando los cebadores de acuerdo con la reivindicación 3.
5. Vector recombinante, **caracterizado porque** comprende una molécula de ácido nucleico de acuerdo con una cualquiera de las reivindicaciones 2 a 4.
6. Célula eucariota aislada, **caracterizada porque** se transforma con una molécula de ácido nucleico de acuerdo con una cualquiera de las reivindicaciones 2 a 4, un vector de acuerdo con la reivindicación 5 o una cepa neurovirulenta del virus West-Nile de acuerdo con la reivindicación 1.
7. Proteína o péptido, **caracterizado porque** está codificado por una molécula de ácido nucleico de acuerdo con una cualquiera de las reivindicaciones 2 a 4.
8. Uso de la cepa de acuerdo con la reivindicación 1, para la obtención de un anticuerpo policlonal, mediante inmunización de un mamífero no humano.
9. Uso de acuerdo con la reivindicación 8, **caracterizada porque** dicho mamífero no humano es un ratón homocigótico para el alelo *FIV*^r de resistencia a la infección por los virus de la familia de los *Flaviviridae*.
10. Uso de un vector recombinante de acuerdo con la reivindicación 5 o bien de una proteína o de un péptido, de acuerdo con la reivindicación 7, para la obtención de un anticuerpo policlonal o monoclonal mediante inmunización de un mamífero no humano.
11. Modelo de estudio no humano de la sensibilidad/resistencia a la infección por un virus de la familia de los *Flaviviridae*, **caracterizado porque** comprende al menos una cepa neurovirulenta del virus West-Nile de acuerdo con la reivindicación 1.
12. Modelo de estudio de acuerdo con la reivindicación 11, **caracterizado porque** comprende además un ratón homocigótico para el alelo *FIV*^r o *FIV*^s.
13. Procedimiento de detección de una infección por *Flaviviridae*, particularmente por el virus West Nile, **caracterizado porque** comprende:
 - la amplificación de los ARN procedentes de una muestra biológica que se someterá a ensayo usando como cebadores moléculas de ácido nucleico de acuerdo con la reivindicación 3, y

- la secuenciación del producto de amplificación obtenido.

14. Procedimiento de cribado de genes celulares implicados en la resistencia de un mamífero a la infección por un virus de la familia de los *Flaviviridae*, **caracterizado porque** comprende las siguientes etapas:

5

- cultivo de células derivadas de un hospedador (humano o no humano) seleccionado por su resistencia o su sensibilidad a la infección por un *Flaviviridae*,
- infección *in vitro* de dichas células por la cepa neurovirulenta del virus West-Nile de acuerdo con la reivindicación 1,
- detección de genes expresados de manera diferencial en dichas células infectadas.

10

15. Uso del modelo de acuerdo con la reivindicación 11 o la reivindicación 12, para la clasificación de moléculas activas contra una infección vírica debida a un virus de la familia de los *Flaviviridae*.

16. Procedimiento de clasificación de moléculas activas contra una infección por un Flavivirus, **caracterizado por:**

15

- la puesta en contacto de un cultivo de células eucariotas, procedentes de un mamífero sensible a la infección por un *Flaviviridae* con una suspensión viral de la cepa de acuerdo con la reivindicación 1, en presencia o en ausencia de la molécula que se someterá a ensayo, y

20

- detección de la replicación del virus.

17. Variante de la cepa viral de acuerdo con la reivindicación 1, **caracterizada porque** su genoma comprende al menos una mutación solamente en la secuencia nucleotídica correspondiente a la proteína NS5.

25

18. Molécula de ácido nucleico, **caracterizada porque** se selecciona entre el grupo constituido por las secuencias SEQ ID N° 3-11.

19. Procedimiento de detección de una infección por el virus West Nile, **caracterizado porque** comprende:

30

- la amplificación de los ARN o ADN procedentes de una muestra biológica que se someterá a ensayo usando como cebadores moléculas de ácidos nucleicos de acuerdo con la reivindicación 3,
- la detección del producto obtenido.

35

20. Kit de detección de una infección por un flavivirus que comprende al menos: fragmentos de la molécula de ácido nucleico de SEQ ID N° 1 tal como se definen en la reivindicación 3,

- reactivos para visualizar la reacción de hibridación entre dichas moléculas de ácidos nucleicos y el ADN o el ARN presente en la muestra que se someterá a ensayo.

CI	C	MSKKPGGPGKSRVNM LKRGMPRVLSLIGLKRAMLSLIDGKGPIRFVLAL
FLA		MSKKPGGPGKSRVNM LKRGMPRVLSLIGLKRAMLSLIDGKGPIRFVLAL
CI		LAFFRTAIAPTRAVLDWRGYNKQTAMKHLLSFKKELGTLTSAINRRSS
FLA		LAFFRTAIAPTRAVLDWRGYNKQTAMKHLLSFKKELGTLTSAINRRSS
CI	F-M	KQKKRGGKTGIAMIGLIASVGAVTLSNFQGVMMTVNATDVTDVITIPT
FLA		KQKKRGGKTGIAMIGLIASVGAVTLSNFQGVMMTVNATDVTDVITIPT
CI		AAGKNLCIVRAMDVGYMCDOTITYECPVLSAGNDPEDIDCWCTKSAYVVR
FLA		AAGKNLCIVRAMDVGYMCDOTITYECPVLSAGNDPEDIDCWCTKSAYVVR
CI	M	YGRCTKTRHSRRSRRSLTVQTHGESTLANKKGAWMDSTKATRYLVKTESW
FLA		YGRCTKTRHSRRSRRSLTVQTHGESTLANKKGAWMDSTKATRYLVKTESW
CI	E	ILRNPGYALVAAVIGWMLGSENTMQRVVFVLLLVAFAYSFNCLGMSNRD
FLA		ILRNPGYALVAAVIGWMLGSENTMQRVVFVLLLVAFAYSFNCLGMSNRD
CI		FLEGVSGATWVDLVLEGGSCVTIMSKDKPTIDVQGMNMEAANLAEVASYC
FLA		FLEGVSGATWVDLVLEGGSCVTIMSKDKPTIDVQGMNMEAANLAEVASYC
CI		YLATVSDLSKACPTMGEAHNDKRAPAFVCRQGVVDRGWNGCGLF GK
FLA		YLATVSDLSKACPTMGEAHNDKRAPAFVCRQGVVDRGWNGCGLF GK
CI		GSIDTCAKFACSTKAIGRTILKENIKYEVAIFVHGPTTVESHGNYSTQVG
FLA		GSIDTCAKFACSTKAIGRTILKENIKYEVAIFVHGPTTVESHGNYSTQVG
CI		ATQAGRFSITPAAPSYTLKLGEGYGEVTVDCFRSGIDTNAYYVMTVGT KT
FLA		ATQAGRFSITPAAPSYTLKLGEGYGEVTVDCFRSGIDTNAYYVMTVGT KT
CI		FLVHREWFMDLNL PWSSAGSTVWRNRETLMETEEPHATKQSVIALGSGEG
FLA		FLVHREWFMDLNL PWSSAGSTVWRNRETLMETEEPHATKQSVIALGSGEG
CI		ALHQALAGAI PVEFSSNTVRLTSGHLKCRVKNELQLKGT TYGVCSKAFK
FLA		ALHQALAGAI PVEFSSNTVRLTSGHLKCRVKNELQLKGT TYGVCSKAFK
CI		FLGTPADTGNGTVVLELQYTGTDGPKVPISSVASLNDLTPVGRLVTVNP
FLA		FLGTPADTGNGTVVLELQYTGTDGPKVPISSVASLNDLTPVGRLVTVNP
CI		FVSVATANAKVLIELEPPFGDSYIVVGRGEQQINHHWNKSGSSIGKAFTT
FLA		FVSVATANAKVLIELEPPFGDSYIVVGRGEQQINHHWNKSGSSIGKAFTT

FIGURA 1A

CI	TLKGAQRLAALGDTAWDFGSGVGVFTSVGKAVHQVFEGGAFRSLFGGMSWI
FLA	TLKGAQRLAALGDTAWDFGSGVGVFTSVGKAVHQVFEGGAFRSLFGGMSWI
	
	NS1
CI	TQGLLGALLLWMGINARDASIALTFLAVGGVLLFLSVNVHADTGCAIDIS
FLA	TQGLLGALLLWMGINARDASIALTFLAVGGVLLFLSVNVHADTGCAIDIS
	
CI	RQELRCGNGVFIHNDVEAWMDRYKYYFETPOGLAKI IQKAHKEGVCGLRS
FLA	RQELRCGSGVFIHNDVEAWMDRYKYYFETPOGLAKI IQKAHKEGVCGLRS
	
CI	VSRLHQMWELAVKDELNTLLKENGVDLSVVVEKQEGMYKSAPKRLTATTE
FLA	VSRLHQMWELAVKDELNTLLKENGVDLSVVVEKQEGMYKSAPKRLTATTE
	
CI	KLEIGWKANGKSILFAPELANNTFVVGDPETKECPTQNRWNSLEVEDFG
FLA	KLEIGWKANGKSILFAPELANNTFVVGDPETKECPTQNRWNSLEVEDFG
	
CI	FGLTSTRMFLKVRESNTTECDSKIIGTAVKNNLAIHSDLSYWIESRLNDT
FLA	FGLTSTRMFLKVRESNTTECDSKIIGTAVKNNLAIHSDLSYWIESRLNDT
	
CI	WKLERAVLGEVKSCTWPFETTLWGDGILESOLIFVTLAGFRSNHNRPG
FLA	WKLERAVLGEVKSCTWPFETTLWGDGILESOLIFVTLAGFRSNHNRPG
	
CI	YKTQNGQFWDEGRVEIDFDYCPGTTVTLSESCGHRGPATRTTTESGKLIT
FLA	YKTQNGQFWDEGRVEIDFDYCPGTTVTLSESCGHRGPATRTTTESGKLIT
	
	NS2A
CI	DWCCRSTLPFLRYQTDSCGWYGMETRPQRHDEKTLVQSQVNAVYADMD
FLA	DWCCRSTLPFLRYQTDSCGWYGMETRPQRHDEKTLVQSQVNAVYADMD
	
CI	PFQGLLVVFLATQEVLRKRWTAHSMPTALLIALVLVFGGITYTDVLR
FLA	PFQGLLVVFLATQEVLRKRWTAHSMPTALLIALVLVFGGITYTDVLR
	
CI	VILVGAFAESNSGGDVVHLALMATFKIQVFMVASFLKARWTNQENIL
FLA	VILVGAFAESNSGGDVVHLALMATFKIQVFMVASFLKARWTNQENIL
	
CI	MLAAVFFQWYHDAHQILLWEIPDVLNSLAVAWMILRAITTTTTSNVVVP
FLA	MLAAVFFQWYHDAHQILLWEIPDVLNSLAVAWMILRAITTTTTSNVVVP
	
CI	LLALLTPRLRCNLQVYRILLMVGIGSLREKRSAAAKKNGASLLCLAL
FLA	LLALLTPGLRCNLQVYRILLMVGIGSLREKRSAAAKKNGASLLCLAL
	
	NS2B
CI	ASTGLFNPMLAAGLIACDPNRKRGWPATEVMTAVGLMFAIVGGLAELDI
FLA	ASTGLFNPMLAAGLIACDPNRKRGWPATEVMTAVGLMFAIVGGLAELDI
	
CI	DSMAIPMTIAGLMFAAFVISGKSTDMWIERTADISWESDAEITGSSERVD
FLA	DSMAIPMTIAGLMFAAFVISGKSTDMWIERTADISWESDAEITGSSERVD
	
CI	VRDDGENFQLMNOPGAPWKIMMLRMVCLAISAYTPWAILPSVVGFWITL

FIGURA 1B

FLA	VRLDDGDNFQLMNDPGAPWKIMMLRMVCLAISAYTPWAILPSVVGFWITL
	NS3
CI	QYTKRGGVLWDTSPSPKEYKKGDTTTGYYRIMTRGLLGSYQAGAGVMVEGV
FLA	QYTKRGGVLWDTSPSPKEYKKGDTTTGYYRIMTRGLLGSYQAGAGVMVEGV
CI	FHTLWHTTKGAALMSGEGRLDPYWGSEVKEBALCYGGPWKLQHKWNGQDEV
FLA	FHTLWHTTKGAALMSGEGRLDPYWGSEVKEBALCYGGPWKLQHKWNGQDEV
CI	QMIIVVEPGKNVKNVQTKPGVFKTFEGEIGAVTLDFPTGTSGSPFVDRKNGD
FLA	QMIIVVEPGKNVKNVQTKPGVFKTFEGEIGAVTLDFPTGTSGSPFVDRKNGD
CI	VIGLYNGVIMPNGSYISAIVOGERMDEFIPAGFEPEMLRKKQITVLDLN
FLA	VIGLYNGVIMPNGSYISAIVOGERMDEFIPAGFEPEMLRKKQITVLDLN
CI	PGAGKTRRLPQIIKEAINRRLRTAVLAPTRVVAEMAALRGLPIRYQT
FLA	PGAGKTRRLPQIIKEAINRRLRTAVLAPTRVVAEMAALRGLPIRYQT
CI	SAVPREHNGNEIVDMCHATLTHRLMSPHRVPNYNLFVMDERHFTDPASI
FLA	SAVPREHNGNEIVDMCHATLTHRLMSPHRVPNYNLFVMDERHFTDPASI
CI	AARGYISTKVELGEAAAFMTATPPGTSDPFESNSPISDLQTEIPORAW
FLA	AARGYISTKVELGEAAAFMTATPPGTSDPFESNSPISDLQTEIPORAW
CI	NSGYEWITEYTGKTVWVPSVVMGNEIALCLQKAGKVVQLNRKSYETEV
FLA	NSGYEWITEYTGKTVWVPSVVMGNEIALCLQKAGKVVQLNRKSYETEV
CI	PKCKNDWDVFIITTOISEMGANFKASRVISAKSVKPTIITEGEGAVILG
FLA	PKCKNDWDVFIITTOISEMGANFKASRVISAKSVKPTIITEGEGAVILG
CI	EPSAVTAASAAQRRGRIGRNPSQVGDEYCYGGHTNEDDSNFAHWTEARIM
FLA	EPSAVTAASAAQRRGRIGRNPSQVGDEYCYGGHTNEDDSNFAHWTEARIM
CI	PONINMPNGLIAQFYQPEREKVYTHEGEYRLAGEERKNFLELLRTADLPV
FLA	PONINMPNGLIAQFYQPEREKVYTHEGEYRLAGEERKNFLELLRTADLPV
CI	WLAYKVAAAGVSYHRRWCDFGPRNTILEDNNEVEVITKLGGRKILAPR
FLA	WLAYKVAAAGVSYHRRWCDFGPRNTILEDNNEVEVITKLGGRKILAPR
	NS4A
CI	WIDARVYSDHQALKAFKDFASGKRSQIELIEVLGKMPHFHMGKTWEALDT
FLA	WIDARVYSDHQALKAFKDFASGKRSQIELIEVLGKMPHFHMGKTWEALDT
CI	NYVVATAEKGGRAHMALEELPDALQTTIALIALSVMTMGVFFLLNQKNG
FLA	NYVVATAEKGGRAHMALEELPDALQTTIALIALSVMTMGVFFLLNQKNG
CI	IGKIGLGGAVLGVATFFCNMAEVPGTKIAGMLLSLLLMIVLIPEPEKQR

FIGURA 1C

FLA	IGKIGLGGAVLGVATFFCWMAEVPGTKIAGMLLSLLMIVLIPPEKQR
	NS4B
CI FLA	SQTDNQLAVFLICVMTLVSAAVANEMGWLDKTKSDISSLFQGRIEVKENF SQTDNQLAVFLICVMTLVSAAVANEMGWLDKTKSDISSLFQGRIEVKENF
CI FLA	SMGEFLDLAPATAWSLYAVTTAVLTPLLKHLITSQYINTSLTSINVOAS SMGEFLDLAPATAWSLYAVTTAVLTPLLKHLITSQYINTSLTSINVOAS
CI FLA	ALFTLARGFPFVDVGVSALLLAGCWGQVTLTVTVTAATLLFCHYAYMV? ALFTLARGFPFVDVGVSALLLAGCWGQVTLTVTVTAATLLFCHYAYMV?
CI FLA	GWQAEAMRSAQRATAAGIMKNAVVDGIVATDVPELEATTPIIMQKXVGQIM GWQAEAMRSAQRATAAGIMKNAVVDGIVATDVPELEATTPIIMQKXVGQIM
CI FLA	LILVSLAAVVVNPSVKTVREAGILITAAAVTLWENGASSVWNATTAGLC LILVSLAAVVVNPSVKTVREAGILITAAAVTLWENGASSVWNATTAGLC
	NS5
CI FLA	HIMRGWLSCLSIITWTLIKNMEKPGKRGGAAGRTLGEVWKERLNQMTKE HIMRGWLSCLSIITWTLIKNMEKPGKRGGAAGRTLGEVWKERLNQMTKE
CI FLA	EFTRYRKEATIEVDOSAAXHAKKEGNVTGGXSVSRGTAKLRWLVERFLE EFTRYRKEATIEVDOSAAXHAKKEGNVTGGXSVSRGTAKLRWLVERFLE
CI FLA	FVGKVIDLGGGRGWCYYMATQKRVQGEVRYTKGGPGHEEPQLVQSYGWN FVGKVIDLGGGRGWCYYMATQKRVQGEVRYTKGGPGHEEPQLVQSYGWN
CI FLA	IIVTKSGVDVFFYPSECCDTLLCDIGESSSSAEVEEHRTIRVLEMVEDWL IIVTKSGVDVFFYPSECCDTLLCDIGESSSSAEVEEHRTIRVLEMVEDWL
CI FLA	HRGPREFCVKVLCFYMPKVIEMKELLQRRYGGGLVANPLSRNSTHENYVW HRGPREFCVKVLCFYMPKVIEMKELLQRRYGGGLVANPLSRNSTHENYVW
CI FLA	SRASGNVHVSNNMTSQVLLGRMEKRTWNGPQYIEDVNLGSGTRAVGKPLL SRASGNVHVSNNMTSQVLLGRMEKRTWNGPQYIEDVNLGSGTRAVGKPLL
CI FLA	NSDTSKINNRIRALAREYSSSTWKKDENHFPYATWNYHGSYDVNFTGSASSL NSDTSKINNRIRALAREYSSSTWKKDENHFPYATWNYHGSYDVNFTGSASSL
CI FLA	VNGVRLLSKPNWDTITNVTTHAMTDTTFGQQRVFKEKVDKAPPEFEGA VNGVRLLSKPNWDTITNVTTHAMTDTTFGQQRVFKEKVDKAPPEFEGA
CI FLA	KYVLNETTNWLWAFLAAREKRPKMCREEFIKVNNSNAALGAMFEEQNQWR KYVLNETTNWLWAFLAAREKRPKMCREEFIKVNNSNAALGAMFEEQNQWR
CI	SAREAVEDPKFWEMVDEEREHLRGECHTCIYNMMGKREKKPGEFGKXKG

FIGURA 1D

FLA	SAREAVEDPKFWEMVDEEREAEHLRGECHTCIYNMMGKREKKPGFPGKAKG
CI FLA	SRAIWFMWLGARFLEFEALGFLNEDHWLGRKNSGGGVEGLGLQKLGYYLR SRAIWFMWLGARFLEFEALGFLNEDHWLGRKNSGGGVEGLGLQKLGYYLR
CI FLA	EVGTRPGGKIYADDTAGWDTRITRADLENEAKVLELLDGEHRRRLARAIIE EVGTRPGGKIYADDTAGWDTRITRADLENEAKVLELLDGEHRRRLARAIIE
CI FLA	LTYRHKVVKVMRPAADGRTVMDVISREDQRGSGQVVTYALNTFTNLAVQL LTYRHKVVKVMRPAADGRTVMDVISREDQRGSGQVVTYALNTFTNLAVQL
CI FLA	VRMEGEGVIGPDDVEKLTGKGKPKVRTWL FENGEERLSRMAVSGDDCVV VRMEGEGVIGPDDVEKLTGKGKPKVRTWL FENGEERLSRMAVSGDDCVV
CI FLA	KPLDDRFATSLHFLNAMSKVRKDIQEWKPSTGWYDWQQVPFCSNHFTELI KPLDDRFATSLHFLNAMSKVRKDIQEWKPSTGWYDWQQVPFCSNHFTELI
CI FLA	MKDGRTLVVPCRGQDELVGRARISPGAGWNVRDTACLAKSYAQMWLLLYF MKDGRTLVVPCRGQDELVGRARISPGAGWNVRDTACLAKSYAQMWLLLYF
CI FLA	HRDLRLMANAICSAVPVNWVPTGRTTWSIHAGGEWNTTEDMLEVWNRVW HRDLRLMANAICSAVPVNWVPTGRTTWSIHAGGEWNTTEDMLEVWNRVW
CI FLA	IEENEWMEDKTPVEKWSQVPYSGKREDIWCGLIGTRARATWAENIQVAI IEENEWMEDKTPVEKWSQVPYSGKREDIWCGLIGTRARATWAENIQVAI
CI FLA	NQVRAIIGDEKYVDYMSSLKRYEDTTLVEDTVL NQVRAIIGDEKYVDYMSSLKRYEDTTLVEDTVL

FIGURA 1E

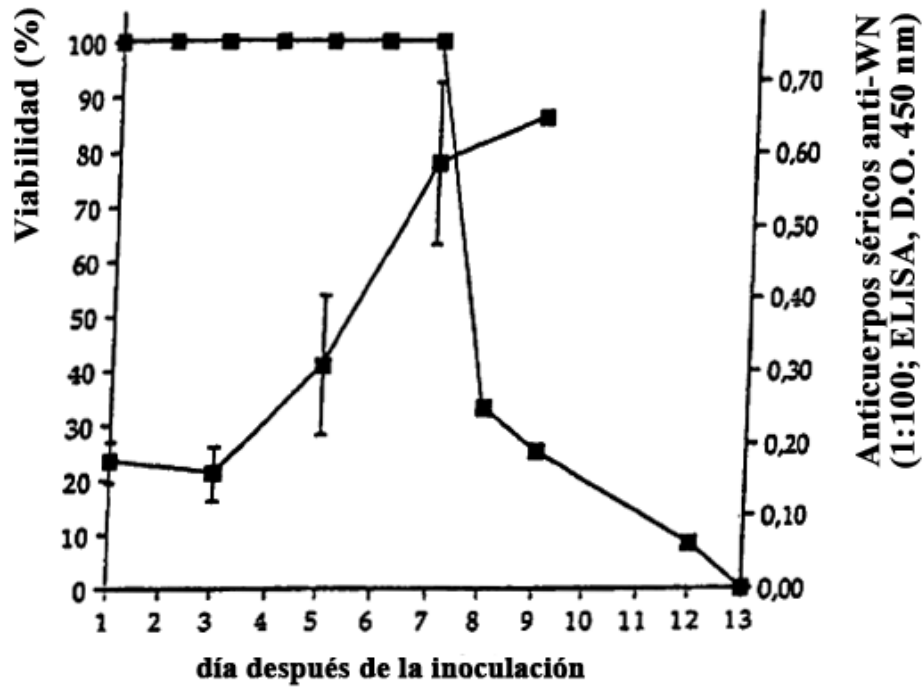


FIGURA 2

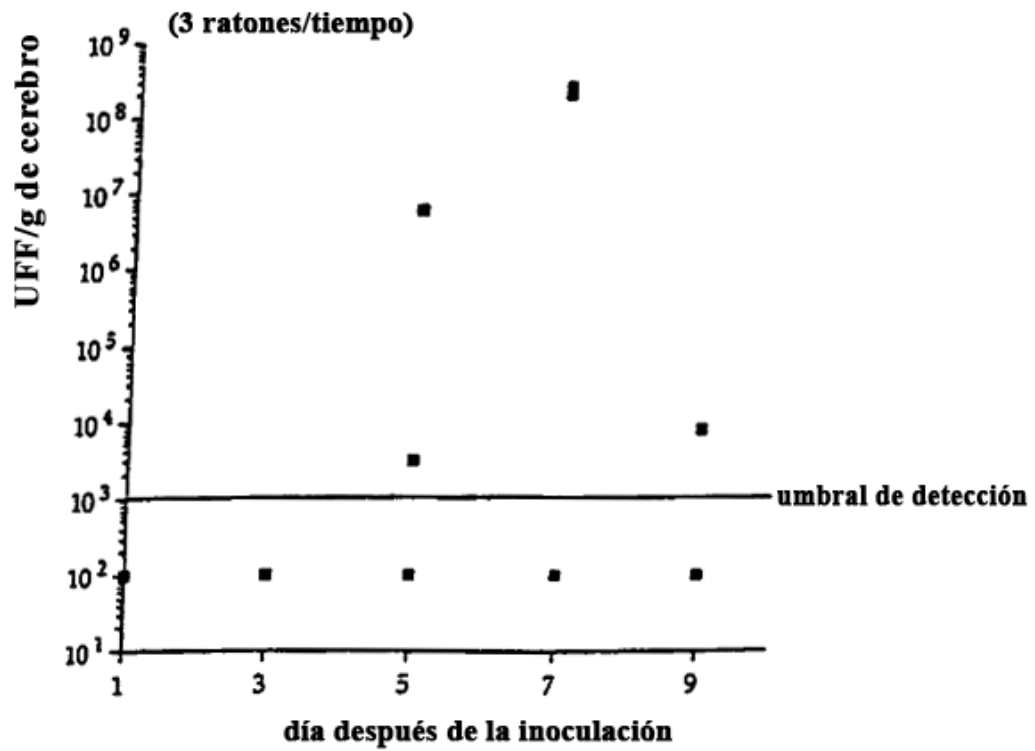


FIGURA 3

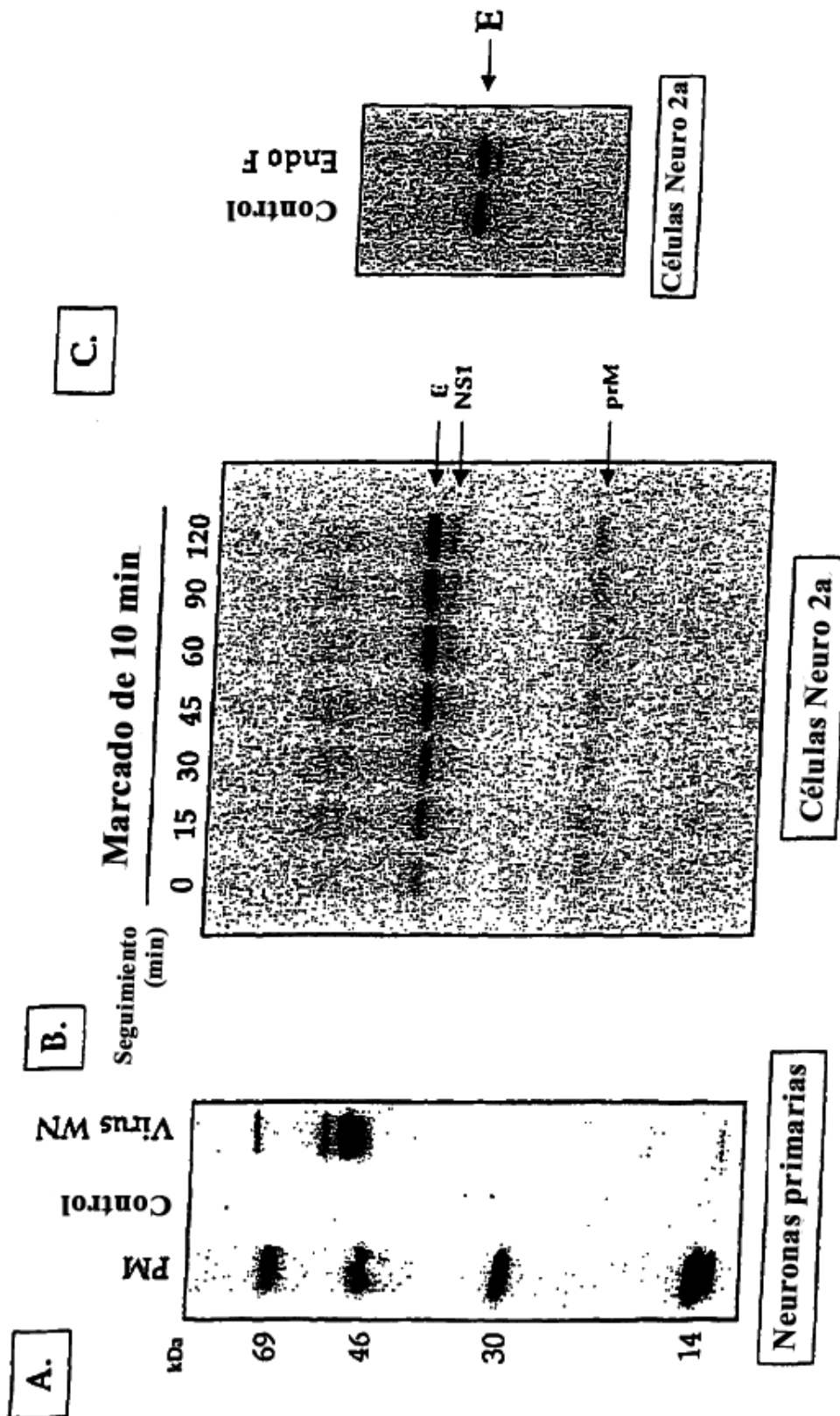
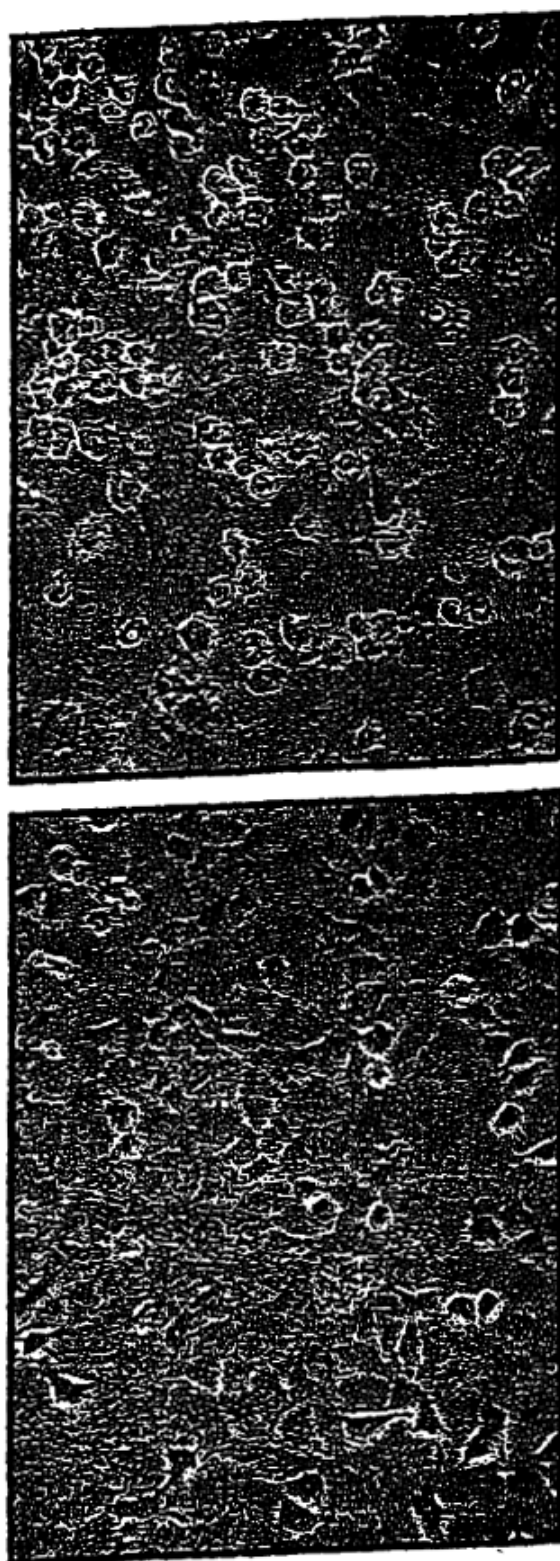


FIGURA 4



Control

**Cepa IS-98-ST1 (m.i. de 4)
24 h de infección**

FIGURA 5

Ratones

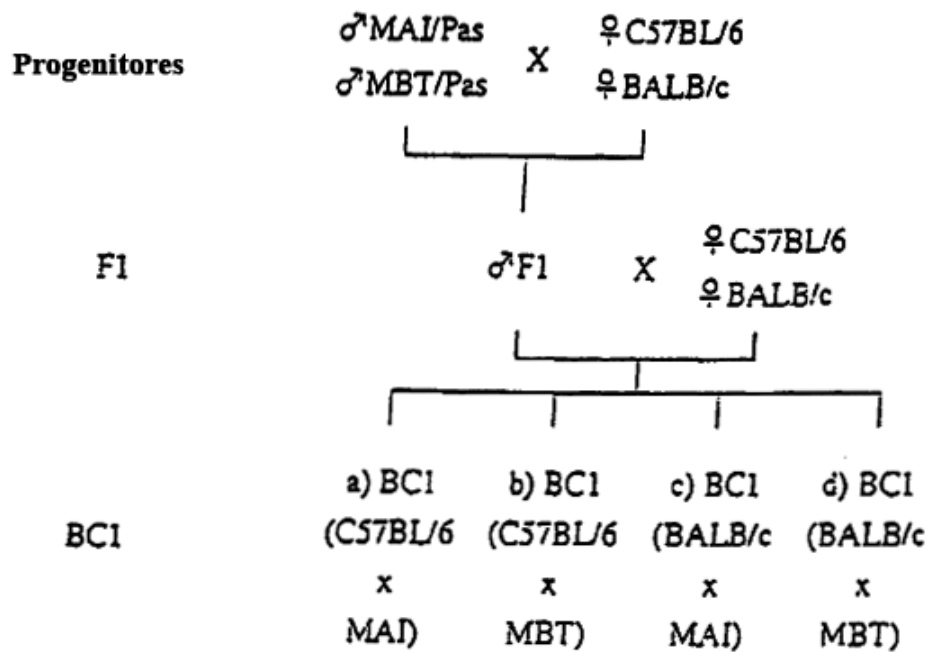
Líneas consanguíneas de laboratorio:

BALB/c, C57BL/6, DDK, 129, C4H y DBA/1
sensibles a la infección por el virus WN

Ratones silvestres:

SEG/Pas (*Mus spretus*), MAI/Pas, MBT /Pas (*Mus m. musculus*)
resistentes a la infección por el virus WN

Generación de ratones de primer retrocruzamiento (BC1)



Virus

Inyección del virus West Nile (WN)

Ratones de 5 semanas de edad

Observación de ratones durante los 14 días después de la infección

Genotipado de los alelos *Flv*

marcadores que flanquean al locus *Flv* en el cromosoma 5 del ratón

FIGURA 6

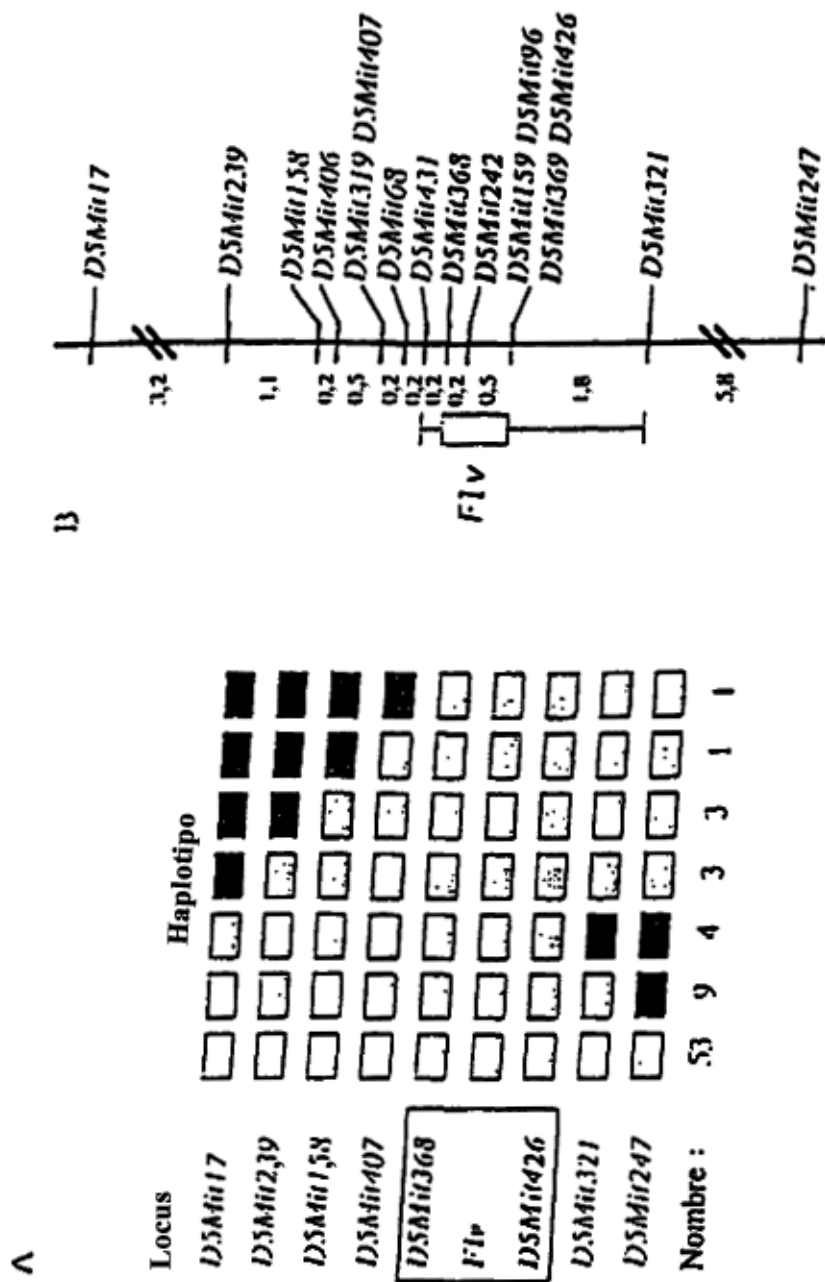


FIGURA 7

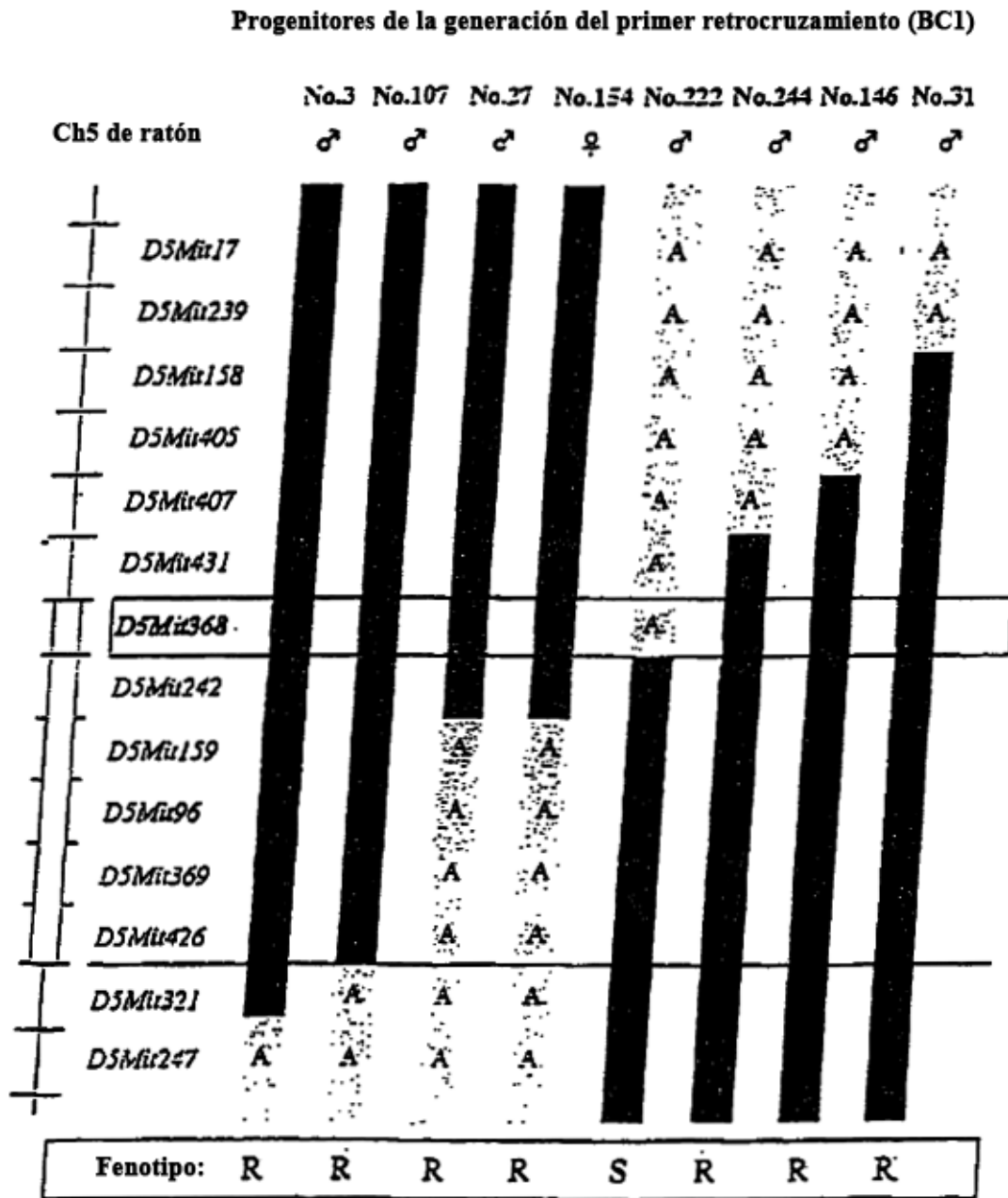


FIGURA 8

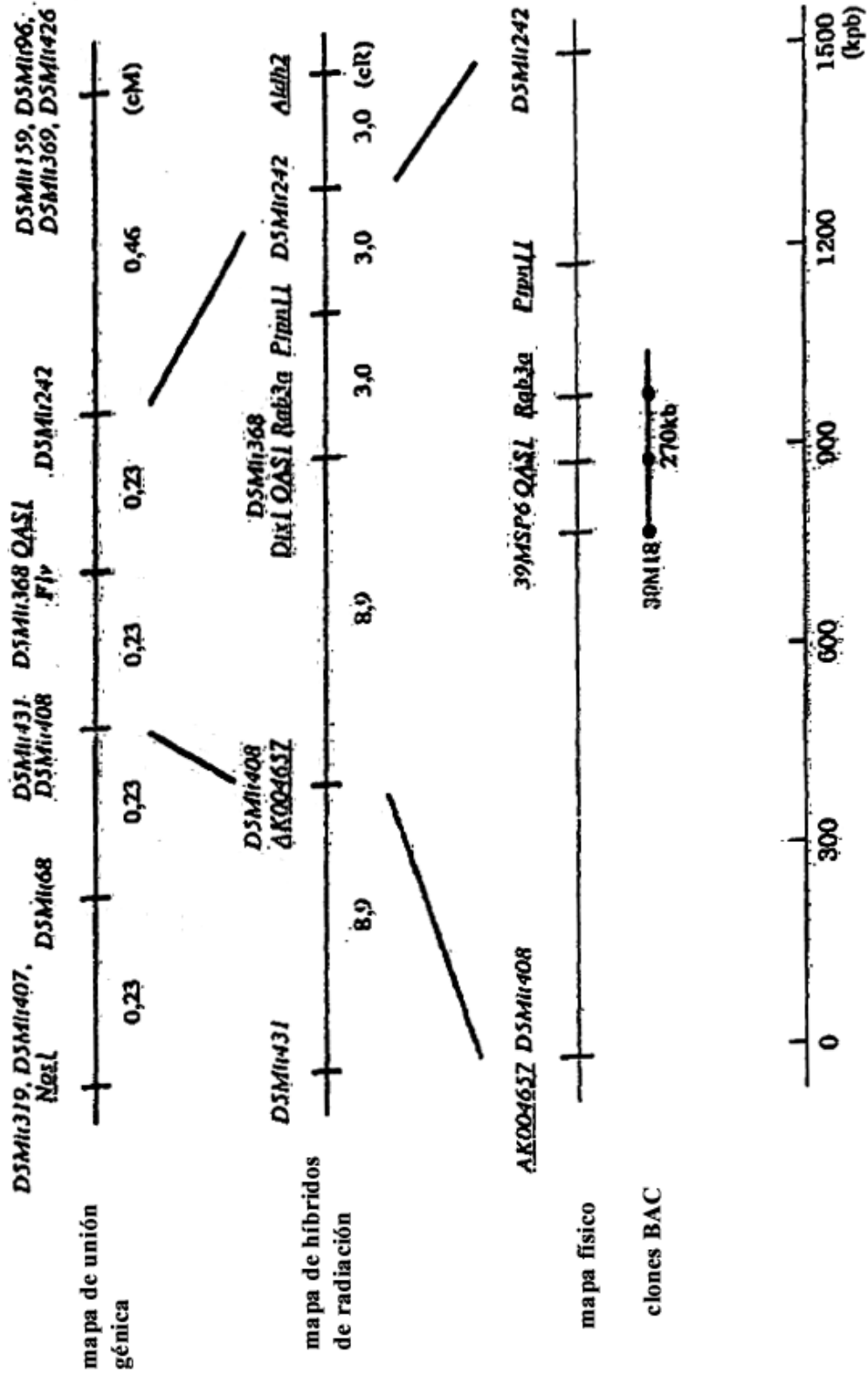


FIGURA 9

Fenotipo <i>Flv</i>*	Supervivientes	Muertos	Total
Resistente (<i>Flv</i>/<i>Flv</i>)	108	0	108 (55%)
Sensible (<i>Flv</i>/<i>Flv</i>)	21	74	95 (45%)
Total	129 (66%)	74 (34%)	203

* un alelo *Flvr* es suficiente para otorgar la resistencia

Figura 10