

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 657 705**

51 Int. Cl.:

C07K 14/005 (2006.01)

C12Q 1/70 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **23.06.2010** E 10006541 (6)

97 Fecha y número de publicación de la concesión europea: **08.11.2017** EP 2399928

54 Título: **Secuencias de virus TT específicas y moléculas de ADN quiméricas de células hospedadoras de virus TT para uso en el diagnóstico, prevención y tratamiento del cáncer y la autoinmunidad**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
06.03.2018

73 Titular/es:

**DEUTSCHES KREBSFORSCHUNGSZENTRUM
(100.0%)
Im Neuenheimer Feld 280
69120 Heidelberg, DE**

72 Inventor/es:

**ZUR HAUSEN, HARALD y
DE VILLIERS, ETHEL-MICHELE**

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 657 705 T3

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

DESCRIPCIÓN

Secuencias de virus TT específicas y moléculas de ADN quiméricas de células hospedadoras de virus TT para uso en el diagnóstico, prevención y tratamiento del cáncer y la autoinmunidad

CAMPO DE LA INVENCION

- 5 La presente invención se refiere a nuevas secuencias de cadena sencilla de una secuencia de virus TT específica y ADN de una célula hospedadora que son capaces de replicarse de forma autónoma para uso como vectores de genes y cromosomas artificiales.

ANTECEDENTES DE LA TECNOLOGÍA

- 10 Desde su descubrimiento en 1997 por Okamoto y sus colegas, se ha observado que los virus TT (VTT) están difundidos ampliamente en todas las poblaciones humanas, en los animales domésticos y en los primates del viejo mundo (1,2). Se ha identificado una gran cantidad de tipos y pseudotipos en los seres humanos, lo que apunta a una notable heterogeneidad de esta familia de virus que ahora se designa como una nueva familia de virus, *Anelloviridae* (3). El ADN vírico se puede encontrar en el suero de casi todos los seres humanos y algunos informes incluso documentan un ADN de este tipo en niños recién nacidos y en sangre del cordón, lo que sugiere una transmisión pre-natal de estos agentes (4,5). A pesar de la presencia generalizada de estos virus, una investigación intensa realizada durante más de 10 años no ha podido demostrar un papel patógeno de tales infecciones en una enfermedad humana.

- 20 Los virus TT no se han replicado con éxito en células de cultivo de tejidos humanos, aunque existen indicios de que la replicación se puede conseguir en células humanas de origen epitelial o hematopoyético. Recientemente, ciclos replicativos de virus del grupo herpes (virus de Epstein-Barr) parecen ejercer un efecto potenciador sobre la amplificación de los genomas de VTT latentes o transfectados (6). Además, los virus TT con frecuencia ponen de manifiesto reordenamientos intramoleculares que conducen a genomas de ADN subvirales en parte defectuosos y con nuevos marcos de lectura abiertos. Se replican de forma autónoma durante períodos prolongados de tiempo en los tejidos infectados (7). Estos ADNs subvirales se encuentran en materiales de biopsia humanos normales y malignos. 25 Una región muy conservada del VTT y cebadores anidados para detectar un fragmento de 71 nucleótidos de dicha región altamente conservada, se describen en Peng, Y.H. et al.; Archives of Virology, 147(1):21-41, 2002.

- 30 Durante los últimos años, se han compilado algunos datos que son indicativos de una asociación de la infección con virus TT con tumores malignos humanos. Una tasa elevada de carga de virus TT se ha observado en una biopsia de bazo de un paciente con linfoma de Hodgkin (24 genotipos individuales de VTT) (8). Del mismo modo, otros informes describen una mayor tasa de prevalencia de VTT en el cáncer colorrectal y de esófago y en cánceres malignos hematopoyéticos, en comparación con tejido no tumoral del mismo paciente o de otros pacientes (9,10). Sin embargo, la ubicuidad de estas infecciones hizo que una interpretación de estos resultados fuera bastante difícil y no permitió relacionar estas observaciones con el desarrollo de tumores.

- 35 El documento WO 00/46407 describe cebadores o sondas útiles para la detección de VTT y vectores basados en VTT de longitud completa. El documento WO 2007/130519 describe micromatrices de hibridación para detectar virus. El documento WO 03/023027 describe secuencias de ADN que comprenden el genoma del virus TT asociadas a tumores y el documento WO 2008/138619 describe un poli(ácido nucleico) de VTT reordenado, asociado con la leucemia infantil.

COMPENDIO DE LA INVENCION

- 40 El problema técnico subyacente de la presente invención es identificar secuencias de VTT específicas que podrían estar claramente asociadas con enfermedades como el cáncer o enfermedades autoinmunes y, por tanto, proporcionar medios para el diagnóstico y la terapia.

- 45 La solución a dicho problema técnico se consigue proporcionando las realizaciones caracterizadas en las reivindicaciones. Observaciones recientes de la persistencia de una región conservada de 71 bases (HCR), presente con solo una variación menor en todos los aislados de VTT caracterizados hasta la fecha, en una mayor cantidad en cánceres humanos y líneas celulares inmortalizadas, fueron algo sorprendentes de cara a la larga no permisibilidad regular de las mismas células para el ADN de VTT transfectado. Aún más sorprendente ha sido la asociación de secuencias de ADN de células hospedadoras en una forma aparentemente de cadena sencilla con la VTT-HCR. La frecuencia y la regularidad de la persistencia de este material genético obviamente extracromosómico en un gran número de 50 líneas celulares de cáncer, así como en biopsias de tejido cerebral afectado de pacientes con esclerosis múltiple, ha conducido a la siguiente hipótesis: genes de una célula hospedadora, ya sea modificados en el proceso recombinatorio o alterados por nuevas secuencias reguladoras de VTT, desempeñan un papel importante en la carcinogénesis humana y también en algunas reacciones autoinmunes. Incluso pueden reemplazar o, en cierto sentido, corresponderse funcionalmente con infecciones por retrovirus en células de roedor y de pollo.

- 55 ***Un nuevo papel para los virus TT en el cáncer humano y la autoinmunidad***

La sorprendente observación de ADN de células hospedadoras asociado a una forma aparentemente de cadena sencilla de la HCR del virus TT, es la base para la siguiente conclusión: todavía no se ha demostrado que secuencias víricas de TT estén integradas en el ADN celular de cadena doble, continuando dentro de los cromosomas de la célula hospedadora. Por lo tanto, el hallazgo opuesto de ADN de la célula hospedadora, relacionado en un estado de cadena sencilla a la HCR de VTT, debería tener una importancia biológica. Los datos actuales indican su persistencia a largo plazo como episomas en líneas celulares de cáncer humano, lo que apunta a un papel de esa persistencia en la proliferación celular. Dos aspectos parecen requerir una consideración específica: un posible papel de estos recombinantes en el cáncer y en la autoinmunidad.

Una posibilidad es la integración aleatoria de secuencias de la célula hospedadora en episomas de VTT. Esto puede suceder después de un desplazamiento de la cadena en el curso de una replicación de ADN aberrante o después de una transcripción inversa del ARN celular. En el caso de una integración aleatoria, una mayor cantidad de recombinantes debe ser inocua e inofensiva para las células que son portadoras de estos recombinantes. Una propiedad que favorece el crecimiento de los transcritos de la HCR de VTT, así como la integración y la transcripción de los genes de la célula hospedadora que estimulan el crecimiento, su modificación en el proceso de integración o su alteración por la HCR de VTT, darán lugar, sin embargo, a consecuencias proliferativas. Estos episomas deben adquirir propiedades inmortalizantes y, bajo ciertas condiciones, propiedades transformadoras. En combinación con modificaciones adicionales del genoma de la célula hospedadora, pueden dirigir el crecimiento maligno. Este modo de acción revela una semejanza distante con la inserción de oncogenes celulares en genomas retrovirales.

El concepto de oncogén de VTT

Las consideraciones anteriores se resumen en la Fig. 4. Obviamente, la recombinación entre la región reguladora de VTT y los ácidos nucleicos celulares debe ser un proceso relativamente frecuente, ya que tales recombinantes se encuentran en la mayoría de las líneas celulares analizadas hasta el momento. También debe contribuir a la proliferación celular, de lo contrario la persistencia regular de tales moléculas, en parte durante décadas de proliferación continua, sería difícil de explicar. Se supone que este tipo de recombinación es un proceso aleatorio, que implica diferentes tipos de genes celulares. La función codificante de la HCR de VTT y/o la captación de genes que dirige la proliferación celular, o que bloquea la función de antagonistas de proliferación, o la inhibición de la diferenciación celular, deberían dar lugar a una acumulación de células que contienen estos tipos de recombinantes. Se prevé que esto, en combinación con eventos mutacionales o recombinantes adicionales de las células que albergan tales recombinantes de ácido nucleico de células hospedadoras de VTT, proporciona una ventaja selectiva para las células que son portadoras de este tipo de episomas. La presencia de estos últimos representaría un factor de riesgo principal para la conversión a cancerígeno. En este sentido, esas recombinaciones deben tener una importancia general para diferentes tipos de cánceres humanos, aunque se espera para tipos de cánceres individuales un cierto grado de especificidad para un conjunto limitado de genes.

Las implicaciones de este modelo son profundas. Alcanzan desde la prevención del cáncer, la detección temprana, hasta la terapia del cáncer. El papel importante de las infecciones con VTT y de la persistencia de HCR de VTT se destaca en la información disponible. La prevención de estas infecciones debe reducir el riesgo de desarrollar los recombinantes descritos. El diagnóstico de recombinantes específicos contribuirá probablemente a la evaluación del riesgo de cáncer. Se esperan profundas implicaciones en la terapia del cáncer: la HCR de VTT surge como el principal determinante de la persistencia y el mantenimiento de los episomas de una sola hebra. Puesto que esta región parece formar parte de un marco de lectura abierto, debería ser vulnerable frente a ARNs o ADNs pequeños de interferencia. Por lo tanto, ofrece una diana adecuada para futuras deliberaciones terapéuticas.

Otros dos aspectos merecen un examen: ciertos paralelos que parecen existir en la carcinogénesis retroviral en roedor y pollo, y el uso de sistemas de vectores basados en VTT de replicación autónoma para la terapia génica. La mutagénesis por inserción, la captación y la modificación de genes que estimulan el crecimiento celular que los convierte en oncogenes, se han analizado frecuentemente en sistemas animales. Esto, por lo tanto, todavía no se ha descrito para los cánceres humanos. ¿Los virus TT reemplazan este nicho en las células de humanos y otros primates? ¿Compite el VTT con éxito con infecciones por retrovirus para hacerse cargo de su papel en especies específicas? La persistencia episómica del ADN de cadena sencilla surge, sin embargo, como una notable diferencia en la carcinogénesis inducida por retrovirus.

Las moléculas de ADN subvirales con replicación autónoma de aproximadamente 400 bases, originarias de VTT, se han descrito antes (11). Es tentador especular que ellas o recombinantes específicos de células hospedadoras de VTT puedan representar sistemas de vectores óptimos para futuros enfoques en la terapia génica y para la construcción de cromosomas artificiales.

El concepto de autoinmunidad del ADN de células hospedadoras de VTT recombinantes

La existencia de recombinantes de ácido nucleico de células hospedadoras de VTT también permite un enfoque novedoso de aspectos de enfermedades autoinmunes y otras enfermedades crónicas (potencialmente, incluso afecciones como la arteriosclerosis y la enfermedad de Alzheimer). La modificación o la alteración de proteínas celulares pueden proceder de eventos por inserción de genes celulares en el ADN de una sola cadena o sobre las diferentes HCRs, ejercida por elementos de VTT (Fig. 5). Podrían proporcionar una explicación conveniente para las reaccio-

nes autoinmunes, incluso para las locales, al igual que en la esclerosis múltiple (MS, del inglés "multiple sclerosis") o la enfermedad de Crohn. En los dos últimos casos, en particular, la reactivación de otras infecciones locales (potencialmente virus de tipo herpes) proporcionaría un estímulo para una amplificación local y la actividad génica de los respectivos recombinantes de ácido nucleico de células hospedadoras de VTT. En la MS, esto podría explicar episodios recurrentes de la progresión de la enfermedad. Un modelo del concepto de autoinmunidad se representa en la Figura 5.

Del mismo modo, se han identificado moléculas de virus TT reordenadas, de 719, 642, y 621 bases que se replican de manera autónoma después de la transfección de líneas celulares específicas. La composición de sus ADNs y el proceder de genotipos de VTT completos específicos, se muestra en la Figura 6. Aquí los reordenamientos dan como resultado nuevos marcos de lectura abiertos, en parte con epítomos relacionados con los de la diabetes juvenil y la artritis reumatoide.

Conclusión

Los modelos de la presente invención para un papel de los recombinantes de ácido nucleico de células hospedadoras de VTT, se basan en la demostración de las moléculas químicas de cadena sencilla entre la HCR de VTT y el ADN de la célula hospedadora y moléculas de VTT reordenadas con replicación autónoma de pesos moleculares sustancialmente reducidos. Ambos, el concepto de oncogén de VTT y el concepto de autoinmunidad de VTT, ofrecerán sin duda nuevos enfoques para la prevención, el diagnóstico y, en particular, para la terapia de estas afecciones y mejorarán el pronóstico de los pacientes respectivos.

Por lo tanto, la presente invención se refiere a un poli(ácido nucleico) de virus TT, en el que el poli(ácido nucleico) de virus TT consiste en

(a) una secuencia de nucleótidos de 71 bases (HCR) que se muestra en la Figura 6;

o (b) una secuencia de nucleótidos que es el complemento de la secuencia de nucleótidos de (a).

La presente invención proporciona un vector de expresión que comprende un poli(ácido nucleico) de virus TT (HCR) de la invención, ligado funcionalmente a elementos de control de la transcripción y traducción procariontas o eucariotas, así como una célula hospedadora que comprende el vector de expresión.

DESCRIPCIÓN DETALLADA DE LA INVENCION

Definiciones

A menos que se defina de otro modo, todos los términos técnicos y científicos usados en este documento tienen el mismo significado que entienden comúnmente los expertos normales en la técnica a la que pertenece la invención. Aunque se puede emplear cualquier método y material, similar o equivalente a los descritos en el presente documento, en la práctica o al someter a ensayo la presente invención, se describen los métodos y materiales preferidos. Para los fines de la presente invención, se definen a continuación los siguientes términos.

La expresión "*muestra biológica*" tal y como se emplea en esta memoria se refiere a una muestra que se puede extraer, sin tratar, tratada, diluida o concentrada a partir de un animal. Muestra biológica se refiere a cualquier muestra biológica (tejido o fluido) que contiene un poli(ácido nucleico) de VTT de la invención y se refiere más particularmente a muestras de suero sanguíneo, muestras de plasma, muestras de biopsias, muestras de líquido cefalorraquídeo, etc..

Por "*vehículo*" se entiende cualquier sustancia de peso molecular normalmente alto a la que una sustancia no inmunogénica o poco inmunogénica (por ejemplo, un hapteno) está ligada de forma natural o artificial para mejorar su inmunogenicidad.

El término "*diagnosis*" se emplea en esta memoria en su sentido más amplio para incluir la detección de un antígeno reactivo con una molécula que se une a un antígeno de una subinmunoglobulina. También se incluye dentro de su alcance, el análisis de mecanismos de trastornos. De acuerdo con ello, el término "diagnosis" incluye el uso de anticuerpos monoclonales para fines de investigación, como herramientas para detectar y comprender los mecanismos asociados con una enfermedad o afección de interés. También incluye el uso diagnóstico del poli(ácido nucleico) de VTT de la invención para la detección de ARN homólogo o complementario transcrito a partir de tales moléculas.

El término "*inmunogenicidad*" se emplea en esta memoria en su sentido más amplio para incluir la propiedad de provocar una respuesta inmune dentro de un organismo. La inmunogenicidad depende normalmente en parte del tamaño de la sustancia en cuestión, y en parte de en qué grado es no similar a las moléculas hospedadoras. Se considera generalmente que las proteínas altamente conservadas tienden a tener una inmunogenicidad más bien baja.

El término "*paciente*" se refiere a pacientes de origen humano o de otro mamífero, e incluye a cualquier individuo que se desea examinar o tratar usando los métodos de la invención. Sin embargo, se entenderá que "*paciente*" no

implica que los síntomas estén presentes. Mamíferos adecuados que están comprendidos dentro del alcance de la invención incluyen, pero no se limitan a, primates, animales de granja (por ejemplo, ovejas, vacas, caballos, burros, cerdos), animales de ensayo de laboratorio (por ejemplo, conejos, ratones, ratas, cobayas, hámsteres), animales de compañía (por ejemplo, gatos, perros) y animales silvestres cautivos (por ejemplo, zorros, ciervos, dingos).

- 5 Por "*vehículo farmacéuticamente aceptable*" se entiende una carga sólida o líquida, diluyente o sustancia encapsulante que se puede utilizar de manera segura en cualquier tipo de administración.

La expresión "*enfermedad o afección relacionada*" se emplea en esta memoria para referirse a una enfermedad o una afección que se relaciona anatómica, fisiológica, patológica y/o sintomáticamente con una enfermedad o afección de referencia. Por ejemplo, las enfermedades o afecciones pueden estar relacionadas entre sí mediante una afección a localizaciones anatómicas similares (por ejemplo, que afectan al mismo órgano o parte del cuerpo), que afectan a diferentes órganos o partes del cuerpo con función fisiológica similar (por ejemplo, el esófago, el duodeno y el colon que se basan en el peristaltismo para mover los alimentos desde un extremo del canal alimentario a otro), por tener patologías similares o solapantes (por ejemplo, lesión o ruptura de un tejido, apoptosis, necrosis) o por tener síntomas similares o solapantes (es decir, respuesta alérgica, inflamación, linfocitosis). Así, por ejemplo, un antígeno asociado con una colitis ulcerosa también se puede asociar con una perforación del colon, ya que esa enfermedad afecta al mismo órgano (es decir, el colon).

El término "*tratar*" se emplea en esta memoria en su sentido más amplio para incluir tanto un tratamiento terapéutico como profiláctico (es decir, preventivo), diseñado para mejorar la enfermedad o la afección.

20 El término "*episoma*" se emplea en esta memoria para referirse a una porción de material genético que puede existir de forma independiente del cuerpo principal del material genético (cromosoma) algunas veces o de manera continua y replicarse de manera autónoma, mientras que otras veces es capaz de integrarse en el cromosoma. Ejemplos de episomas incluyen secuencias de inserción, transposones y el VTT de la invención.

LEYENDAS DE LAS FIGURAS

25 Figura 1: amplificación con PCR de un fragmento de 71 bases que contiene la región de VTT altamente conservada (HCR) en 4 líneas celulares diferentes, L1236 (línea de linfoma de Hodgkin negativa para EBV), HSB-2 (línea de leucemia linfoblástica aguda), KR e IGL (líneas celulares de melanoma) y ADN de placenta.

Figura 2: ADN enrollado que permanece en el material sobrenadante de las células L1236 después de una precipitación y recogida del ADN de peso molecular elevado y digestión con ARNasa

Dos bandas son visibles en la región entre las bandas de 4,3 y 6,6 bases.

30 Figura 3: PCR larga dirigida hacia afuera usando cebadores de la región HCR de VTT de 71 bases en ADN de HSB-2

Dos bandas son visibles en regiones correspondientes a 4,5 a 7 kb. Además, las bandas surgen en la región correspondiente a 0,4 a 0,7 kb.

Figura 4: Representación esquemática del concepto de oncogén de VTT

35 La parte izquierda representa la organización genómica de los genomas de VTT de tipo silvestre. La parte derecha prevé la integración del ADN de la célula hospedadora en los plásmidos de cadena sencilla.

Figura 5: Representación esquemática del concepto de autoinmunidad del ADN de células hospedadoras de VTT

Los genes modificados de las células hospedadoras deben codificar epítomos antigénicos inmunorreactivos.

Figura 6: Amplificación con PCR de la HCR de 71 bases a partir del ADN de 4 líneas celulares diferentes

40 Las flechas apuntan a los dos sitios con variaciones en las secuencias de nucleótidos.

Figura 7:

45 **(A)** El ADN con replicación autónoma de 719 bases de VTT (derecha) y la secuencia de VTT completa a partir de la cual se obtiene. La composición de nucleótidos de ambas moléculas se encuentra en la Figura 11A + B.

(B) El ADN con replicación autónoma de 621 bases de VTT (derecha) y la secuencia de ADN completa a partir de la cual se obtiene. La composición de nucleótidos de ambas moléculas se encuentra en la Figura 12A + B.

(C) El ADN con replicación autónoma de 642 bases de VTT (derecha) y la secuencia de ADN completa a partir de la cual se obtiene. La composición de nucleótidos de ambas moléculas se encuentra en la Figura 13A +

B.

Figura 8: Tres secuencias ejemplares de ADN quimérico de VTT / célula hospedadora truncada procedentes de biopsias cerebrales de pacientes con esclerosis múltiple

5 (A) Secuencias celulares quiméricas obtenidas a partir del cromosoma 1 con algunas homologías con secuencias de priones y de tumor de Wilms y el extremo 3' del 3 seudogén de leucemia linfocítica mieloide 3 (MLL3). Secuencia de ADN humano del clon RP11-14N7 en el cromosoma 1. Contiene un extremo 3' de un seudogén de leucemia linfocítica mieloide 3 o un linaje mixto (MLL3), un seudogén del receptor de 7 hélices transmembranales, el extremo 5' de un gen novedoso.

10 (B) Secuencias celulares quiméricas obtenidas a partir del cromosoma 16. Homologías con el factor de transcripción 3 (TF 3C), firmas proteicas para receptores de quimiocinas y receptor de leucotrienos B4.

(C) Secuencias celulares quiméricas obtenidas a partir del cromosoma 10, secuencia truncada de la miosina, reactividad descrita en pacientes con esclerosis múltiple y aquellos con artritis reumatoide (la secuencia contiene ambos cebadores completos delantero y posterior).

15 Figura 9: Tres secuencias ejemplares de ADN quimérico de VTT / célula hospedadora truncada procedentes de líneas celulares obtenidas a partir de pacientes con la enfermedad de Hodgkin o leucemia

(A) Secuencias del cromosoma 1 con parte de transgelina 2, el gen IGSF9 para miembros de la superfamilia de inmunoglobulinas 9, el gen SLAMF9.

20 (B) Secuencias de proteínas traducidas con homología sustancial con los oncogenes v-myb (oncogén vírico de la mieloblastosis aviar), pero también con c-myb. Esta secuencia se amplificó con el cebador directo en ambos extremos.

(C) Obtenida a partir del cromosoma 10. Homología elevada con la proteína "Deleted in malignant 1 Protein" (DMBT), un gen supresor tumoral identificado. Esta secuencia se amplificó con el cebador directo en ambos extremos.

25 Figura 10: Secuencias de cebadores utilizadas en las reacciones descritas en los Ejemplos, obtenidas a partir de la HCR de 71 bases.

Figura 11:

(A) Secuencia completa de VTT a partir de la cual se ha obtenido el ADN de 719 bases que se replica autónomamente.

(B) Secuencia completa del ADN de TTV de 719 bases que se replica autónomamente.

30 Figura 12:

(A) Secuencia completa de VTT (tth25) a partir de la cual se ha obtenido el ADN de 621 bases que se replica autónomamente.

(B) Secuencia completa del ADN de TTV de 621 bases que se replica autónomamente.

Figura 13:

35 (A) Secuencia completa de VTT (ttrh215) a partir de la cual se ha obtenido el ADN de 642 bases que se replica autónomamente.

(B) Secuencia completa del ADN de TTV de 642 bases que se replica autónomamente.

Figura 14: Marcos de lectura abierta (ORFs) que se encuentran dentro de la secuencia de nucleótidos de 71 nt

40 zyb2.1.pep, zyb9.1.pep y zkb69.1.pep están empezando en el primer triplete, zyb2.3.pep, zyb9.3.pep, zkb5.3.pep y zkb69.3.pep están empezando a partir del tercer triplete. Esta región se transcribe activamente.

Figura 15: Digestión de ADN monocatenario mediante la nucleasa de judía de mungo (MBN). Los carriles 2 y 3 muestran que el ADN amplificado puede ser digerido mediante un pretratamiento con MBN. Los carriles 5 y 6 muestran que ADN plasmídico tratado previamente de la misma manera, no es digerido por MBN.

45 La presente invención proporciona un poli(ácido nucleico) del virus TT que consiste en una secuencia de nucleótidos mostrada en la Figura 6 como se define en la reivindicación 1.

El experto en la técnica puede determinar fácilmente qué secuencias de ácidos nucleicos están relacionadas con la secuencia de nucleótidos de la Figura 6 o qué fragmentos todavía son capaces de replicarse de forma autónoma

mediante el uso de ensayos convencionales o los ensayos descritos en los ejemplos, a continuación.

La presente invención se refiere más específicamente a un poli(ácido nucleico) del virus TT que tiene (a) una secuencia de nucleótidos mostrada en la Figura 6, o (b) el complemento de (a).

5 El término "poli(ácido nucleico)" se refiere a una secuencia de ácido nucleico de cadena sencilla o de cadena doble. Un poli(ácido nucleico) puede consistir en desoxirribonucleótidos o ribonucleótidos, análogos de nucleótidos o nucleótidos modificados, o se puede haber adaptado para fines terapéuticos. El poli(ácido nucleico) del virus TT es un ADN de cadena sencilla.

10 El poli(ácido nucleico) del virus TT de la invención puede estar presente como un episoma extracromosómico, puede estar integrado en el genoma del hospedador y/o puede estar ligado a un ADN de la célula hospedadora, por ejemplo, un ADN que comprende un gen de la célula hospedadora que estimula el crecimiento o que contiene genes truncados de la célula hospedadora con inmunogenicidad alterada.

Preferiblemente, el polinucleótido del virus TT de la invención comprende una secuencia de nucleótidos que se selecciona a partir del grupo de secuencias de nucleótidos que se muestra en las Figuras 8 y 9.

15 Las secuencias de polinucleótidos de VTT que son similares a las secuencias tal y como se muestran en la Figura 6, se pueden caracterizar y aislar de acuerdo con cualquiera de los métodos conocidos en la técnica, tales como la amplificación por medio de cebadores específicos de secuencia, la hibridación con sondas específicas de secuencia en condiciones más o menos rigurosas, la determinación de la secuencia de la información genética de VTT, etc.

20 La presente invención se refiere a un vector de expresión recombinante que comprende un poli(ácido nucleico) de VTT de la invención tal y como se ha definido anteriormente, ligado funcionalmente a elementos de control de la transcripción y traducción procariotas o eucariotas.

El término "vector" puede comprender un plásmido, un cósmido, un cromosoma artificial, un fago o un virus. Particularmente útiles para el desarrollo de vacunas pueden ser moléculas recombinantes de virus TT, BCG o vectores adenovíricos, así como virus recombinantes de la viruela aviar.

25 La expresión "expresado de forma recombinante" empleada en el contexto de la presente invención, se refiere al hecho de que los polipéptidos de la presente invención se producen mediante métodos de expresión recombinantes ya sea en procariotas o eucariotas inferiores o superiores, como se describe más detalladamente a continuación.

30 La expresión "eucariota inferior" se refiere a células hospedadoras tales como levaduras, hongos y similares. Los eucariotas inferiores son por lo general (pero no necesariamente) unicelulares. Eucariotas inferiores preferidos son levaduras, particularmente especies dentro de *Saccharomyces*, *Schizosaccharomyces*, *Kluiveromyces*, *Pichia* (por ejemplo, *Pichia pastoris*), *Hansenula* (por ejemplo, *Hansenula polymorpha*), *Schwaniomyces*, *Schizosaccharomyces*, *Yarrowia*, *Zygosaccharomyces* y similares. *Saccharomyces cerevisiae*, *S. carlsbergensis* y *K. lactis* son los hospedadores de levadura empleados más comúnmente, y son hospedadores fúngicos convenientes.

35 La expresión "eucariota superior" se refiere a células hospedadoras obtenidas a partir de animales superiores, tales como mamíferos, reptiles, insectos y similares. Células hospedadoras eucariotas superiores preferidas actualmente se obtienen a partir de células de hámster chino (por ejemplo, CHO), mono (por ejemplo, COS y células Vero), riñón de hámster bebé (BHK), riñón de cerdo (PK15), células de riñón de conejo 13 (RK13), la línea celular de osteosarcoma humano 143 B, la línea celular humana HeLa y líneas celulares de hepatoma humano tales como Hep G2, y líneas celulares de insecto (por ejemplo, *Spodoptera frugiperda*). Las células hospedadoras se pueden proporcionar en cultivos en suspensión o en matraz, cultivos de tejidos, cultivos de órganos y similares.

40 El término "procariotas" se refiere a hospedadores tales como *E. coli*, *Lactobacillus*, *Lactococcus*, *Salmonella*, *Streptococcus*, *Bacillus subtilis* o *Streptomyces*. También estos hospedadores se contemplan dentro de la presente invención.

45 La expresión "célula hospedadora" se refiere a células que se pueden haber utilizado o se han usado como receptores para un vector recombinante u otro polinucleótido de transferencia, e incluyen la progenie de la célula original que se ha transfectado. Se entiende que la progenie de una única célula parental puede que no sea necesariamente completamente idéntica en la morfología o en el complemento de ADN genómico o total, a la parental original, debido a una mutación o recombinación natural, accidental o deliberada.

50 El término "replicón" es cualquier elemento genético, por ejemplo, un plásmido, un cromosoma, un virus, un cósmido, etc., que se comporta como a una unidad autónoma de replicación de polinucleótidos dentro de una célula, es decir, es capaz de una replicación bajo su propio control.

El término "vector" es un replicón que comprende además secuencias que proporcionan una replicación y/o una expresión de un marco de lectura abierto deseado.

La expresión "elemento de control" se refiere a secuencias de polinucleótidos que son necesarias para efectuar la

expresión de secuencias codificantes a las que están ligadas. La naturaleza de tales secuencias de control difiere dependiendo del organismo hospedador; en procariotas, tales secuencias de control incluyen generalmente un promotor, un sitio de unión al ribosoma, sitios de corte y empalme y terminadores; en eucariotas, generalmente, tales secuencias de control incluyen promotores, sitios de corte y empalme, terminadores y, en algunos casos, potenciadores. La expresión "elementos de control" se entiende que incluye, como mínimo, todos los componentes cuya presencia es necesaria para la expresión, y también pueden incluir componentes adicionales cuya presencia es ventajosa, por ejemplo, secuencias líder que gobiernan la secreción.

El término "promotor" es una secuencia de nucleótidos que se compone de secuencias de consenso que permiten la unión de la ARN polimerasa con el molde de ADN, de manera que esa producción de ARNm se inicia en el sitio normal de iniciación de la transcripción para los genes estructurales adyacentes.

La expresión "ligado funcionalmente" se refiere a una yuxtaposición en la que los componentes descritos de este modo están en una relación que les permite actuar de la manera pretendida. Una secuencia de control "ligada funcionalmente" a una secuencia codificante, está ligada de tal manera que la expresión de la secuencia codificante se consigue en condiciones compatibles con las secuencias de control.

El segmento del ADN de VTT que codifica la secuencia deseada insertada en la secuencia del vector, se puede fijar a una secuencia señal. Dicha secuencia señal puede ser la que proviene de una fuente no VTT, pero las estructuras artificiales particularmente preferidas de acuerdo con la presente invención, contienen secuencias señal que aparecen en el genoma de VTT antes de los respectivos puntos de inicio de las proteínas.

Los eucariotas superiores se pueden transformar con vectores, o pueden estar infectados con un virus recombinante, por ejemplo, un virus vaccinia recombinante. Las técnicas y los vectores para la inserción de ADN extraño en virus vaccinia son bien conocidos en la técnica, y utilizan, por ejemplo, la recombinación homóloga. Una amplia variedad de secuencias promotoras víricas, posiblemente secuencias terminadoras y secuencias de adición de poli(A), posiblemente secuencias potenciadoras y posiblemente secuencias de amplificación, todas necesarias para la expresión en mamíferos, están disponibles en la técnica. El virus vaccinia es particularmente preferido ya que el virus vaccinia detiene la expresión de las proteínas de células hospedadoras. Para una vacunación de los seres humanos, la viruela aviar y el virus Ankara modificado (AMV) son vectores particularmente útiles.

También se conocen vectores de transferencia de expresión de insectos obtenidos a partir del baculovirus, virus de la poliedrosis nuclear *Autographa californica* (AcNPV), que es un vector de expresión vírica independiente auxiliar. Los vectores de expresión obtenidos a partir de este sistema normalmente utilizan el potente promotor del gen de polihedrina vírica para dirigir la expresión de genes heterólogos. Diferentes vectores así como métodos para la introducción de ADN heterólogo en el sitio deseado de baculovirus están disponibles para el experto en la técnica, para la expresión de baculovirus. También diferentes señales para una modificación posterior a la traducción, reconocidas por células de insectos, son conocidas en la técnica.

La presente invención también se refiere a una célula hospedadora tal y como se ha definido anteriormente, que comprende un vector recombinante como se ha definido anteriormente.

Los vectores preferidos útiles para la terapia génica son vectores víricos, por ejemplo, adenovirus, virus herpes, virus vaccinia o, más preferiblemente, un virus de ARN tal como un retrovirus. Incluso más preferiblemente, el vector retrovírico es un derivado de un retrovirus murino o aviar. Ejemplos de tales vectores retrovíricos que se pueden emplear en la presente invención son: virus de la leucemia murina de Moloney (MoMuLV), virus del sarcoma murino de Harvey (HaMuSV), virus de tumor mamario murino (MuMTV) y virus del sarcoma de Rous (RSV). Lo más preferiblemente, se emplea un vector retrovírico de primate no humano, tal como el virus de la leucemia de mono gibón (GaLV), proporcionando una gama de hospedadores más amplia, en comparación con los vectores murinos. Ya que los retrovirus recombinantes son defectuosos, se requiere asistencia a fin de producir partículas infecciosas. Esta asistencia se puede proporcionar, por ejemplo, usando líneas de células auxiliares que contienen plásmidos que codifican todos los genes estructurales del retrovirus bajo el control de secuencias reguladoras dentro de la LTR.

Líneas celulares auxiliares adecuadas son bien conocidas por los expertos en la técnica. Dichos vectores pueden contener adicionalmente un gen que codifica un marcador seleccionable, de modo que las células transducidas pueden ser identificadas. Además, los vectores retrovíricos se pueden modificar de tal manera que se convierten en específicos de dianas. Esto se puede lograr, por ejemplo, mediante la inserción de un polinucleótido que codifica un azúcar, un glicolípido o una proteína, preferiblemente un anticuerpo. Los expertos en la técnica conocen métodos adicionales para generar vectores específicos de dianas. Otros vectores y métodos adecuados para la terapia génica *in vitro* o *in vivo* se describen en la bibliografía y son conocidos por las personas expertas en la técnica; véase, por ejemplo, el documento WO 94/29469 o WO 97/00957. Las secuencias de polinucleótidos de VTT de la invención pueden por lo tanto servir ellas mismas como un vector adecuado, compuesto o bien únicamente por las secuencias de ADN químicas de células hospedadoras de VTT. Además, las secuencias de nucleótidos de la invención se pueden utilizar para la construcción de cromosomas artificiales.

Los siguientes ejemplos están destinados a ilustrar, pero no a limitar la invención. Aunque tales Ejemplos son típicos de los que se podrían emplear, otros métodos conocidos por los expertos en la técnica, se pueden utilizar alternativamente.

Ejemplo 1

5 Demostración de la persistencia del ADN de VTT en células procedentes de líneas de cultivo de tejidos obtenidas a partir de tumores malignos

Las líneas celulares obtenidas a partir de tumores malignos poseen una ventaja sobre el material de una biopsia de tumor primario. Comúnmente representan preparaciones puras de células de cáncer, mientras que los materiales primarios generalmente están contaminados por células mesenquimatosas normales, por células del sistema hematopoyético y células epiteliales normales. Por otra parte, un inconveniente de las líneas de cultivo de tejidos puede surgir de la selección de clones específicos que crecen en condiciones de cultivo de tejidos y la adquisición de modificaciones genéticas secundarias durante el curso de un cultivo a largo plazo. Además, los sueros de ternera fetal pueden suponer un riesgo debido a la introducción de virus del ganado que sobreviven después de procedimientos de inactivación del suero (por ejemplo, poliomavirus bovino); véase la Tabla 1 que resume estas ventajas/inconvenientes.

Tabla 1

Análisis de biopsias de tumores primarios frente a líneas celulares establecidas para secuencias relacionadas con el VTT			
Biopsias		Líneas celulares	
Ventaja	Inconveniente	Ventaja	Inconveniente
Materiales auténticos	Contaminadas por mezcla por adición de células normales	Preparaciones puras de células de cáncer	Selección de clones específicos adaptados a las condiciones de cultivo de tejidos
	Búsqueda complicada de secuencias de VTT por la presencia uniforme de VTT en la sangre periférica	Disponibles en cantidades ilimitadas	Cambios genéticos secundarios durante el cultivo a largo plazo
	Disponibilidad limitada		El uso de suero de ternera fetal plantea el riesgo de contaminaciones con virus de ganado

Los intentos de encontrar ADN de VTT en materiales tumorales primarios humanos adolece de un inconveniente: la pluralidad de genotipos de VTT en el material humano (8). Esto hace que sea prácticamente imposible identificar un genotipo específico como un agente etiológico para un tipo de cáncer humano. Por estos motivos, se iniciaron estudios de la persistencia de secuencias de ADN de VTT en células obtenidas a partir de líneas de cultivo de tejidos de cáncer. Hasta ahora, los resultados han sido extremadamente sorprendentes: los cebadores para la PCR usados para descubrir regiones del marco de lectura abierto de VTT han sido del todo infructuosos. Sin embargo, otras combinaciones de cebadores, que descubrían exclusivamente una región reguladora corta, rica en GC del genoma VTT de aproximadamente 71 bases, detectaron esa secuencia en un mayor número de líneas celulares (Figura 1). Esta región reguladora está muy conservada entre diferentes genotipos de VTT y no está presente en la base de datos del genoma humano.

En una primera serie de experimentos, la misma secuencia fue descubierta en una serie de líneas celulares adicionales. Estas incluyen las siguientes líneas:

- MCF7 (línea de cáncer de mama);
- HAK-1, KMH-2, L1236 (todas líneas de linfoma de Hodgkin negativas para el virus Epstein-Barr);
- Y69 (linfoma B negativa para el virus Epstein-Barr)
- HSB-2 (leucemia linfocítica aguda);

- *P3HR-1* (linfoma de Burkitt positiva para el virus Epstein-Barr);
- *BJAB* (linfoma de Burkitt negativa para el virus Epstein-Barr);
- *Ng* (linfoblastos B inmortalizados con EBV de un paciente con esclerosis múltiple).

5 Además de estas 9 líneas positivas, dos líneas celulares de melanoma (IGL y KR, Fig. 1) y ADN de la placenta humana eran negativos en los experimentos iniciales. Curiosamente, después de la eliminación del ADN enrollado de las células L1236 y el tratamiento con ARNasa de la solución restante, además del ADN mitocondrial, se convirtieron en bandas visibles dos bandas débiles de tamaño similar entre las posiciones 4,3-6,6 kb (marcador del tamaño del ADN de cadena doble) en los geles de agarosa (Fig. 2). Un análisis de estas secuencias reveló de nuevo la presencia de la región reguladora de VTT. La nucleasa de semilla de Mungo, que digiere selectivamente el ADN de cadena sencilla, suprimió completamente las bandas que contenían ADN celular procedentes de cuatro biopsias de esclerosis múltiple, en contraste con el ADN de cadena doble de control, lo que subraya la naturaleza de una sola hebra del primero. Estudios similares se están realizando actualmente para aislados de ADN tumoral.

Ejemplo 2

Análisis de secuencias de ADN de VTT quimérico/células hospedadoras truncadas

15 Inicialmente, todos los intentos fallaron para el uso de cebadores con orientación hacia fuera que comenzaban dentro de la región reguladora, con el fin de encontrar ADN del virus TT flanqueante, que rodeaba esta región. Invariablemente, sin embargo, el ADN celular humano se pudo demostrar en los respectivos clones (Fig. 3).

20 Los genes humanos en esos clones y sus disposiciones dentro del ADN episómico de una sola hebra, controlados obviamente por la región de 71 bases de VTT, están siendo analizados actualmente. Los datos disponibles indican una variación sustancial en la captación de genes de células hospedadoras comúnmente truncadas. Su posible conversión en oncogenes estimulantes del crecimiento o en funciones que interfieren con genes supresores tumorales, requiere pruebas funcionales que se están investigando actualmente. Lo mismo explica las secuencias de virus VTT reorganizadas. Algunos de los datos disponibles se presentan en las Figuras 7, 8, 9 y 11 a 13.

Lista de referencias

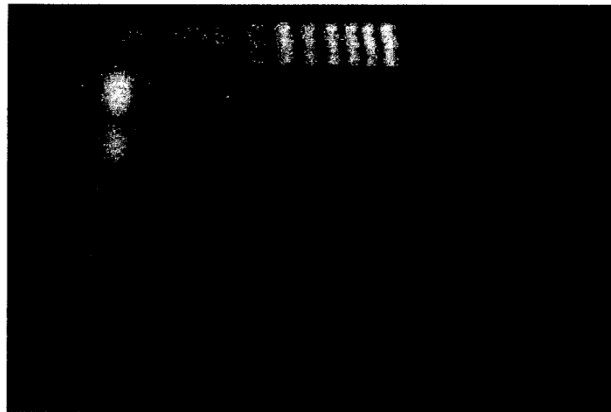
- 25 1. Nishizawa, T., Okamoto, H., Kato, N. et al. A novel DNA virus (VTT) associated with elevated transaminase levels in posttransfusion hepatitis of unknown etiology. *Biochem. Biophys. Res. Commun.* 241: 92-97, 1997.
2. Okamoto, H. TT viruses in animals. Review. En: *TT Viruses, the still Elusive Pathogens* (de Villiers, E.-M., zur Hausen, H. compiladores), *Curr. Topics Microbiol. Immunol.* E.-M. de Villiers, H. zur Hausen (compiladores), Vol. 331: 35-52, 2009.
- 30 3. Biagini, P., Todd, D., Bendinelli, M. et al., Anellovirus. En: *Virus Taxonomy, VIIIth Report of the International Committee for the Taxonomy of Viruses*. London: Elsevier/ Academic Press p. 335-341, 2005.
4. Gerner, P., Oettinger, R., Gerner, W., et al. Mother-to-infant transmission of TT virus: prevalence, extent and mechanism of vertical transmission. *Pediatr. Infect. Dis. J.* 19: 1074-1077, 2000.
- 35 5. Goto, K., Sugiyama, K., Ando, T., et al., Detection rates of TT virus DNA in serum of umbilical cord blood, breast milk and saliva. *Tohoku J., Exp.Med.* 191: 203-207, 2000.
6. Borkosky, S., Whitley, C. y de Villiers, E.-M., resultados no publicados.
7. Leppik, L., Gunst, K., Lehtinen, M., et al., In vivo and in vitro intragenomic rearrangement of TT viruses. *J. Virol.* 81: 9346-9356, 2007.
- 40 8. Jelcic, I., Hotz-Wagenblatt, A., Hunzicker, A., et al. Isolation of multiple TT virus genotypes from spleen biopsy tissue from a Hodgkin's disease patient: genome reorganization and diversity of the hypervariable region. *J. Virol.* 78: 7498-7507, 2004.
9. de Villiers, E.-M., Schmidt, R., Delius, H., et al. Heterogeneity of TT virus related sequences isolated from human tumour biopsy specimens. *J. Mol. Med.* 80: 44-50, 2002.
- 45 10. de Villiers, E.-M., Bulajic, M., Nitsch, C., et al. VTT infection in colorectal cancer tissues and normal mucosa. *Int. J.Cancer* 121: 2109-2112., 2007.

11. de Villiers, E.M., Kimmel, R., Leppik, L. y Gunst, K. Intragenomic rearrangements in TT viruses; a possible role in the pathogenesis of disease. En: TT Viruses, the still Elusive Pathogens (de Villiers, E.-M., zur Hausen, H. compiladores), Curr. Topics Microbiol. Immunol. 331: 91-107, 2009.
- 5 12. Matsukara, M., Shinozuka, K., Zon, G., Mitsuya, H., Reitz, M., Cohen, J., Broder, S. Proc. Natl. Acad. Sci. USA 84(21): 7706-10 (1987).
13. Miller, P., Yano, J., Yano, E., Carroll, C., Jayaram, K., Ts'o, P. (1979) Nonionic nucleic acid analogues. Synthesis and characterization of dideoxyribonucleoside methylphosphonates. Biochemistry 18(23): 5134-43.
14. Nielsen, P., Egholm, M., Berg, R., Buchardt, O. (1991) Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide. Science 254(5037): 1497-500.
- 10 15. Nielsen, P., Egholm, M., Berg, R., Buchardt, O. (1993) Sequence specific inhibition of DNA restriction enzyme cleavage by PNA. Nucleic-Acids-Res. 21(2): 197-200.
16. Asseline, U., Delarue, M., Lancelot, G., Toulme, F., Thuong, N. (1984) Nucleic acid-binding molecules with high affinity and base sequence specificity: intercalating agents covalently linked to oligodeoxynucleotides. Proc. Natl. Acad. Sci. USA 81(11): 3297-301.
- 15 17. Kwok, S., Kellogg, DE., McKinney, N., Spasic, D., Goda, L., Levenson, C., Sninsky, JJ. (1990) Effects of primer-template mismatches on the polymerase chain reaction: human immunodeficiency virus type 1 model studies. Nucleic Acids Res 18: 999-1005.
18. Landgren, U., Kaiser, R., Sanders, J., Hood, L. (1988) A ligase-mediated gene detection technique. Science 241:1077-1080.
- 20 19. Wu, D., Wallace, B. (1989) The ligation amplification reaction (LAR)-amplification of specific DNA sequences using sequential rounds of template-dependent ligation. Genomics 4: 560-569.
20. Walker, GT., Fraiser, MS., Schram, JL., Little, MC., Nadeau, JG., Malinowski, DP. (1992) Strand displacement amplification -- an isothermal, in vitro DNA amplification technique. Nucleic Acids Res 20: 1691-6.
- 25 21. Guatelli, J., Whitfield, K., Kwoh, D., Barringer, K., Richman, D., Gengeras, T. (1990) Isothermal, in vitro amplification of nucleic acids by a multienzyme reaction modeled after retroviral replication. Proc Natl Acad Sci USA 87: 1874-1878.
22. Compton, J. (1991) Nucleic acid sequence-based amplification. Nature 350: 91-92.
- 30 23. Kwoh, DY., Davis, GR., Whitfield, KM., Chappelle, HL., DiMichele, LJ., Gingeras, TR. (1986) Transcription-based amplification system and detection of amplified human immunodeficiency virus type 1 with a bead-based sandwich hybridization format. Proc Natl Acad Sci USA, 86: 1173-7.
24. Duck, P. (1990) Probe amplifier system based on chimeric cycling oligonucleotides. Biotechniques 9: 142-147.
25. Lizardi, PM., Kramer, FR. (1991) Exponential amplification of nucleic acids: new diagnostics using DNA polymerases and RNA replicases. Trends Biotechnol 19: 53-8.
- 35 26. Lomeli, H., Tyagi, S., Pritchard, CG., Lizardi, PM., Kramer, FR. (1989) Quantitative assays based on the use of replicatable hybridization probes. Clin Chem 35: 1826-31.
27. Köhler et al., Nature 256 (1975), 495.
28. Wahl et al., J. Nucl. Med. 24: 316-325 (1983).
- 40 29. Jacobs, KA. et al. (1988) The thermal stability of oligonucleotide duplexes is sequence independent in tetraalkylammonium salt solutions: application to identifying recombinant DNA clones. Nucleic Acids Res 16:4637-50.

30. Zimmermann et al., (1994) Neuron 12, 11-24.
31. Vidal et al.; (1990) EMBO J. 9, 833-840.
32. Mayford et al., (1995), Cell 81, 891-904.
33. Pinkert et al., (1987) Genes & Dev. 1, 268-76).
- 5 34. Nielsen et al., Science 254 (1991), 1497-1500.
35. Fingl et al., The Pharmacological Basis of Therapeutics, Goodman y Gilman, compiladores. Macmillan Publishing Co., New York, p. 1-46 (1975).

REIVINDICACIONES

1. Un vector de expresión que es capaz de replicarse de forma autónoma en una célula hospedadora humana que comprende un poli(ácido nucleico) del virus TT ligado funcionalmente a
- 5 (i) elementos de control de la transcripción y traducción procariota, en donde los elementos de control procariotas comprenden un promotor, un sitio de unión al ribosoma, sitios de corte y empalme y un terminador, o
- (ii) elementos de control de la transcripción y traducción eucariota, en donde los elementos de control eucariotas comprenden un promotor, sitios de corte y empalme y un terminador, y
- los elementos de control de la transcripción y traducción controlan la expresión del poli(ácido nucleico) del virus TT, que consisten en
- 10 (a) una secuencia de nucleótidos seleccionada a partir del grupo consistente en las secuencias que se muestran en la Figura 6 y que codifican uno de los péptidos denominados zyb2.1.pep, zyb9.1.pep, zkb69.1.pep, zyb2.3.pep, zyb9.3.pep, zkb5.3.pep y zkb69.3.pep que se muestran en la Figura 14;
- o
- (b) una secuencia de nucleótidos que es el complemento de la secuencia de nucleótidos de (a).
- 15 2. El vector según la reivindicación 1, en donde el vector comprende un poli(ácido nucleico) del virus TT quimérico del poli(ácido nucleico) del virus TT tal como se ha definido en la reivindicación 1 y un ácido nucleico celular, en donde el poli(ácido nucleico) del virus TT quimérico está ligado funcionalmente a elementos de control de la transcripción y traducción procariotas o eucariotas y el ácido nucleico celular tiene una secuencia de nucleótidos que se selecciona a partir del grupo de secuencias de nucleótidos que consiste en gbDhDi38 que se muestra en la Figura
- 20 8A, gb40.27 que se muestra en la Figura 8B, gb43.30 que se muestra en la Figura 8C, hod11 que se muestra en la Figura 9A, hoht33 que se muestra en la Figura 9B y hoht22 que se muestra en la Figura 9C.
3. El vector de expresión según la reivindicación 1 o 2, en donde el vector se selecciona a partir del grupo que consiste en plásmido, cósmido y cromosoma artificial.
4. Una célula hospedadora que comprende un vector de expresión según una cualquiera de las reivindicaciones 1-3.
- 25



L 1236
HSB-2
Kr
IGL
Placenta
H₂O

Figura 1

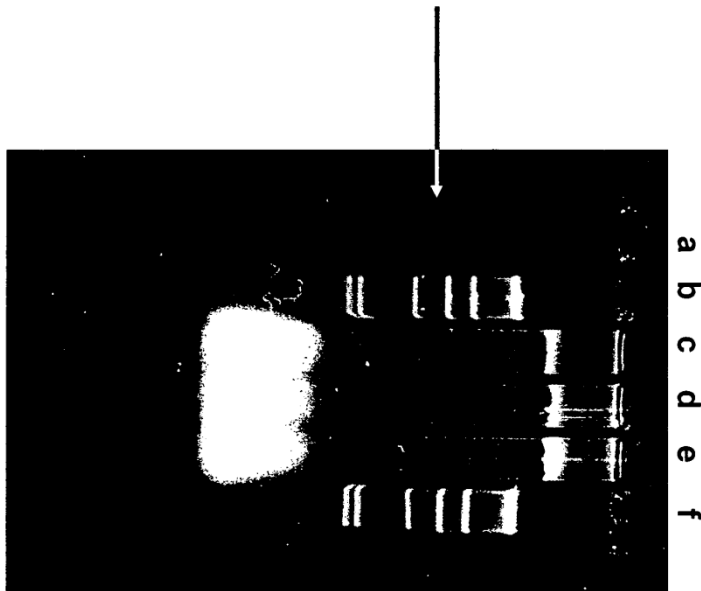


Figura 2

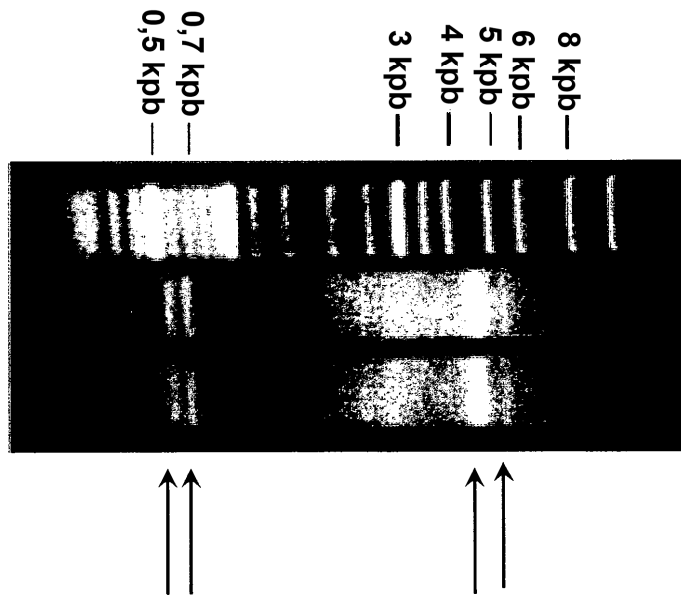


Figura 3

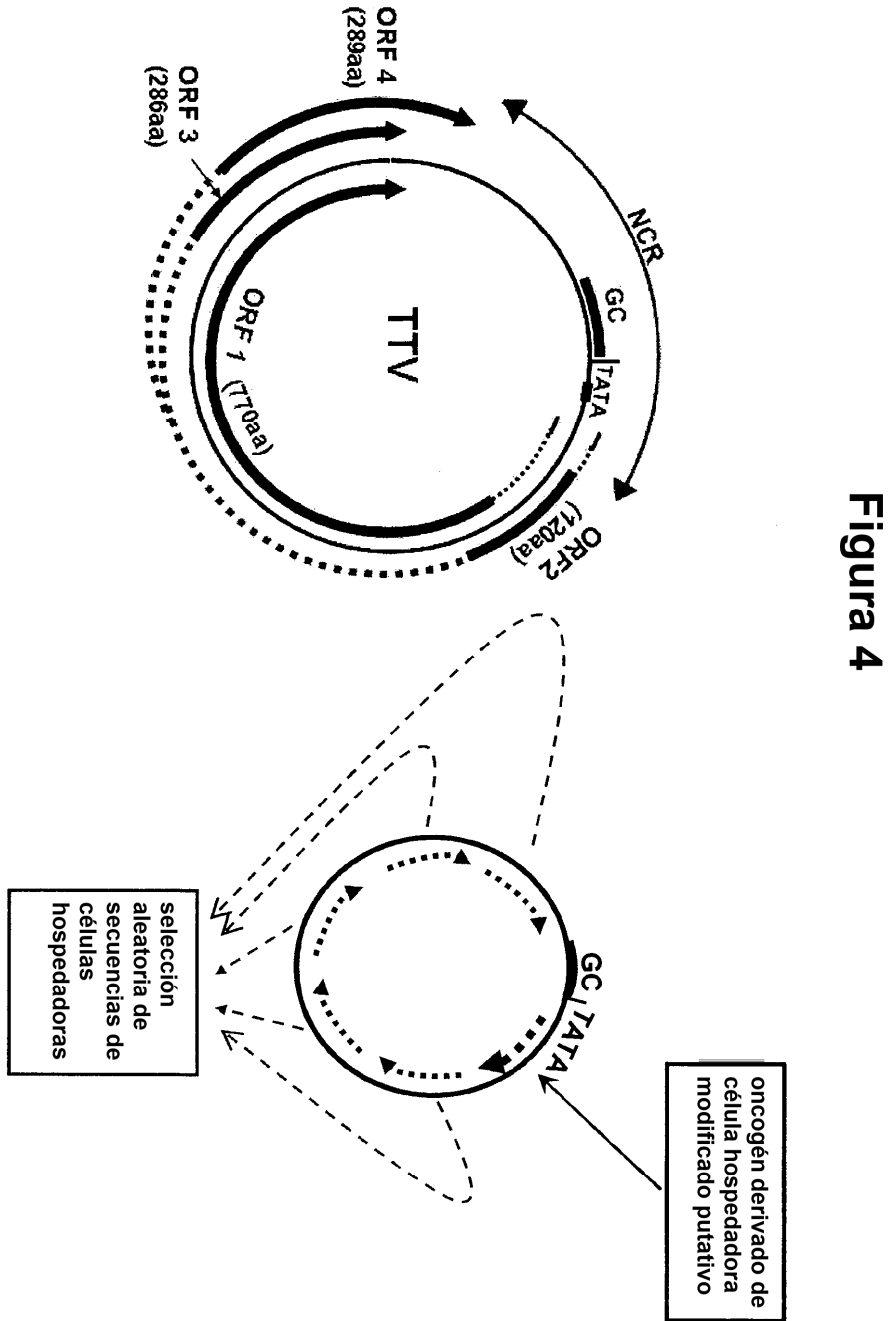


Figura 4

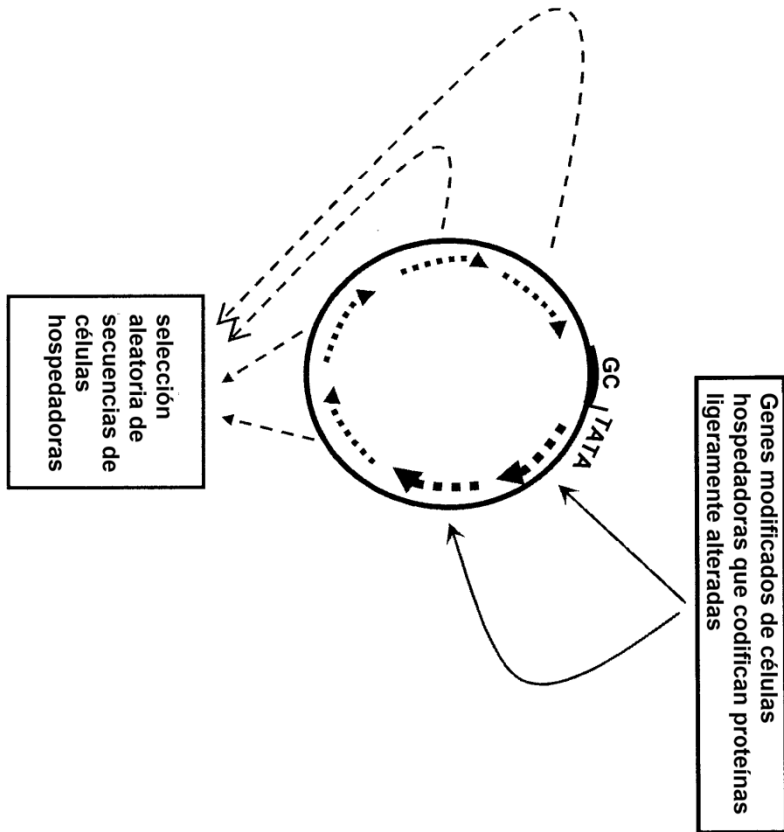
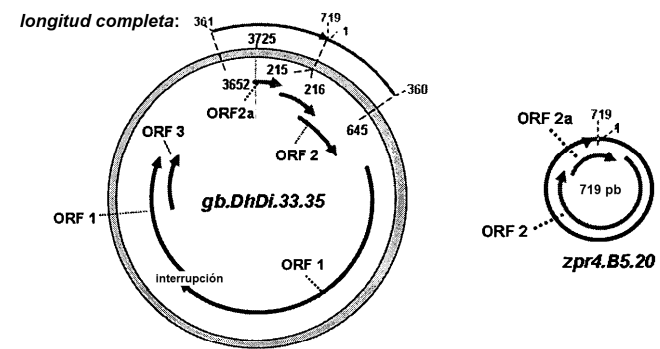


Figura 5

Figura 7:

A



B

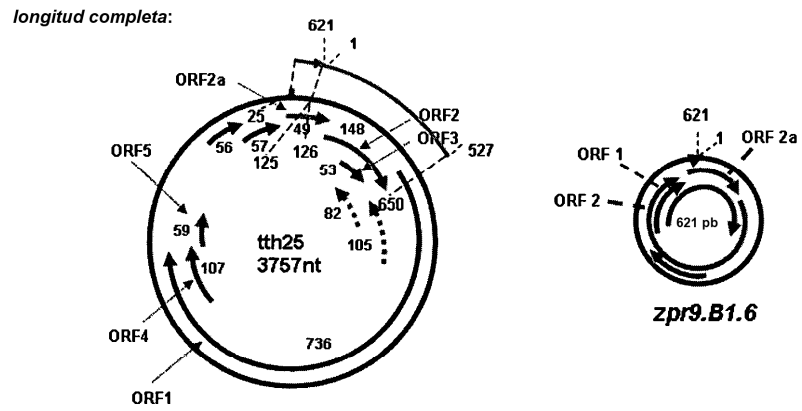


FIGURA 7

C

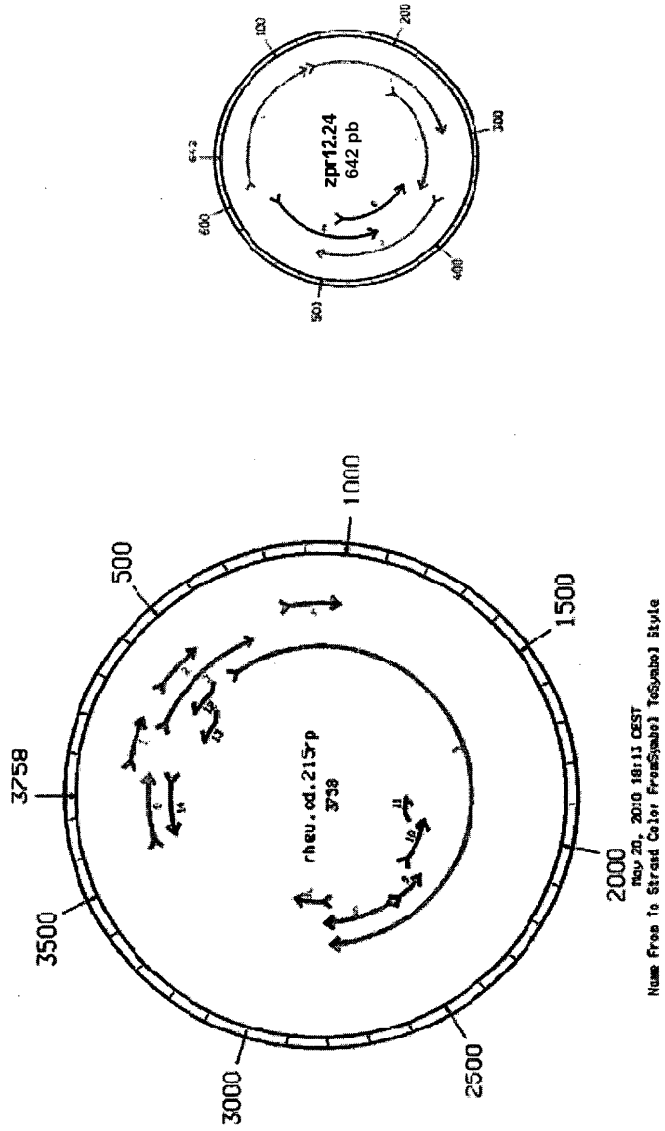
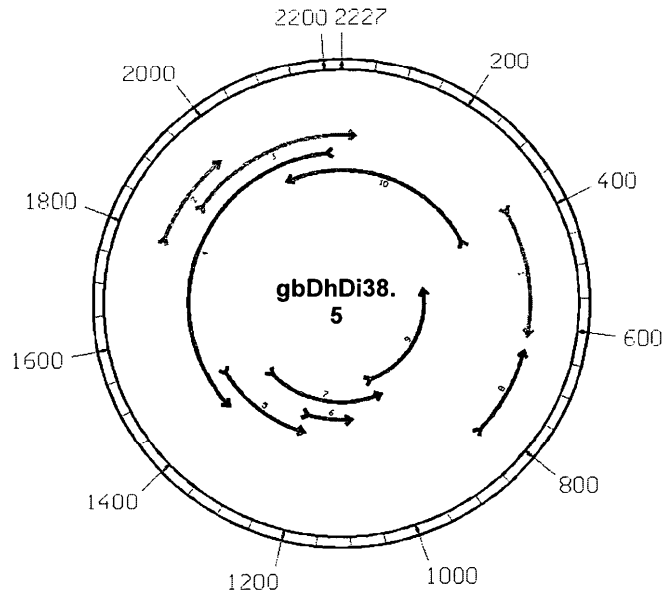


Figura 8

A



gb38.5
gbDhDi38.5

WV13038 Klon6
pcr larga según Templphi
cebador : t3pb1+2 aso. : 65 °C
Longitud : 2227

```

1  CAATTCGTGC ACGGGACTAC AAGGAAAGGG GTTGACCCCC ACCCTCCCCC
51  GCCATGCCCA GGAGGGTGCA GACACAAC TG GGAAGGTGCT AGAGACCCCG
101 GGGGGAGGCT GGGCCAGCAC CAGGCATTGG GGGGCAGGTT CCCGTCTCTA
151 CACCCAGCC CCAGGCGGAC AGCGCGTGCC CCTCCCGCTG CCCACCTGT
201 CACCCACCTG CTGGCCCCGG GCTGTCTCTG CTCTGGGCTC CCCTCCCAGC
251 TGCGTCCCA GCTGCCTCTC CAGGGAGGAG TGACAGCTGG CCTGTGCCAC
301 ACCCTCGAGC CCCCCGGAC TACCCCTCC CTGGGGCAGG ACCCTGCCT
351 GTGGCACAAC CAAGGGCCT GCTGATGGG GCTCATGTGA GCAGTGCCCC
    
```

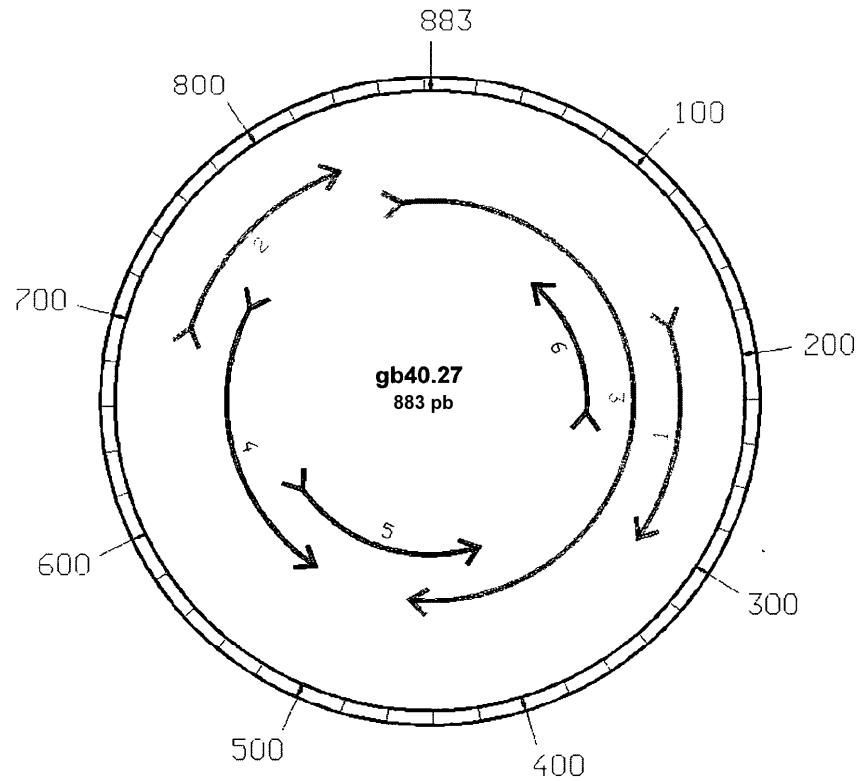
ES 2 657 705 T3

401 AGCTGTGGGT GTGGGTGCTG CCAGCTGCCA CCGCCTTTC CCTGGTTTCC
451 CAGATAGACC CCGACCCACA CTCCGAAGCT GTATCATGAA CGCTGTGGTG
501 GCGGCTGGT GGGGAGCGGG GTTGCCGTCC CACTACCCTC TGAAGCCTC
551 AGCCATGAAG GGCCCTGTG GGCACCTTTT CCCGGCACAC GGTGCTGTGT
601 TTCTCCACTC TTGGGCTCTG CAGTACTTG AGGGGTCAAG TCTATGATCC
651 CACGGGAGGC TGGGCTAATG AGGGGACCAG AGACCTCAGT GCTGTGCAGG
701 GAGTCCTGAA CCACCCTGGT GGAAGGCCA GCCCAACTCC CCAGTCCCTC
751 CGCCAGCTCC CTGTGGTGTG CAGGAGACCT GTGGTCAGGC CTGGAGGAGA
801 AGCTCCTCCT CCCCTCGACA TCCTCCCTGC AGCCCTTGCT CTTACCAGA
851 GCCTCCTGAC TCCCAGGAC CCCAGAGAGG ACTGACCCTC TCCAGCCGAC
901 CTCTGGGCTC AGGACAGCTG GCGGGGCGAG CCACAGGAGC TGCCTGTAGG
951 GAGCAGAGTC AGGACGGGGA CCGAGCCGGA CACCCATTCT GGAAGTGTCT
1001 GCACTCCAG GCAGGGGAAG GACGGCAGTG GGTAGCTGGG AGTGTGGGC
1051 CGAAGATGG CATTGTCAGC CCCTCAGTGG GGACTGGGAG GTAGAGGTGG
1101 GGAGGTCTGT GGAGGAAGGA GAAGAAGGGC CAGTGTCCCG AGTTGGGGGT
1151 GGTGGCAGT GGACGAGCC GACAGGAACA GACCTGAGCT TGGGGAGCTC
1201 CACTCAGAAC GAGGCATCCT TCAGGGTTCT GTGCATACTG GTGTCCCTGG
1251 CTGGGGGCGG GGCCCGAAG TGGAGCCTGG GACTGTGAGG GTGGGGGGGG
1301 TGTGCTGGG TGGGAGGTGG ATGGAGCCCC CCCTCCACCG CCTGGCCGCT
1351 TGGGCTGAAC CTTGGACTTC GGAGCCGAA CAGACATAGG AAATGGCCTA
1401 ACTGCATTTG CGCAGGAACA CCAAATCCCT CGCAGCTGCA CGGGGCTGAG
1451 CCAGGGCCAC GGGCGGGGTC GGCCATCCCA GAGTCCTGAC AGTCCGTGG
1501 TGTATGCCAA GGGCCTGGG CCGCTGACCG AGGGGCGCCT TTCCAGGCC
1551 AGAGGCCCCC ACCCCACCCC AGGAGAGCTG CCCCCTTTC AGTCCCAGA
1601 ACGGAGCCCG GCTGTGGAAT AGTGATGCGG TGAGGTCATG GGGAGGGGGC
1651 CCGCATGACT CATATCCTGG GGTAGGGGAA AGGGAGGAGA CGGAGAAGGG
1701 GCCCAGAGGC CTCCACGTCC TCAGCTCTGC TGGTTCAGAG GCCAGGGGCT
1751 GCGGGGCTT CTCCCAGCA CTGGGTTTTA GGGGAGACAC CAGGAGATGC
1801 TTACTCTGCA TCCCCACTCT GTCCCCAGG CCCCTAGCCA GGGAGAGCTC
1851 AGTCAGAGTG ATCCTCCAGG GGCCAGCTC TGCATGGATG ATGTTCCAG
1901 AGTACACACC TGGGCCTCGT GCCAGGGCCG GCACCGCCGT TGTACGGGCT

ES 2 657 705 T3

1951 ATGGCAAGGC AACAGTCAA TGTTGCCTC ACTAAAGTGA GGCTGCAGCA
2001 CCCTGAAGGG ATCCCTGGAG GGGGACGTGG TCCCCTTGTT CCCAAGCTTG
2051 TCTGCACATG CACGTGGATG TCAAGGGTTC CCGTGTGTGA GCACATGCAT
2101 ATTTGTATGT GCATGGGGTG CGGGCATGTG TGCCTGTGTG GCCGGAGCGT
2151 GGGCTCGTGG AGAATGTGTG TGAGTTGGGT GTGCACCTGC ATGTGCCCCA
2201 GGCCTAGGGA **GTCCCGTGCC CGAATTG**

B



gb40.27

gbDfDg40.27.seq (secuencia celular de gb40 más cerebro de ms adicional)

Longitud : 883

pcr larga según Templphi

cebador: hel 1 + hel 2 aso 68°C (anidada)

vector: pcr2.1 invitrogen

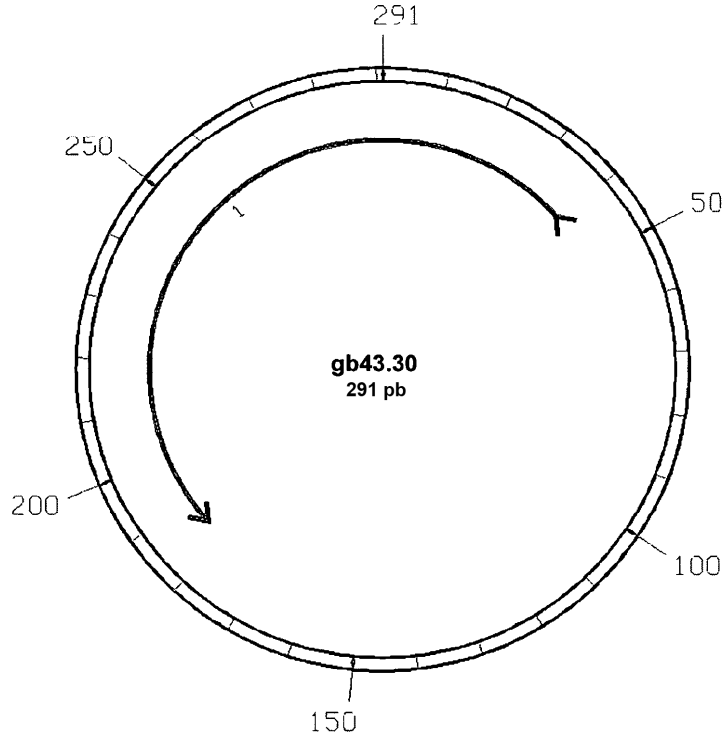
```

1  CGGGACTGGC CGGGCTATGC CCCAGACACA CTCACGTAGG GGTGTCGGC
51  CTGGCAGCCC AGGACCATGG TCTGCAGGGT TTCCTCTCGG CCATTCAGGA
101 CAACCCTAGT CTCCAGGGAA TAGCGCTGGT GTCGCCTATC AGCCGTGAAG
151 GTCTCCTGCA GGAGGAGGCT CTGCGGGATG GGCAGGTGCA ATGGGTGCCT
    
```

ES 2 657 705 T3

201 GGTGTGCAGA GGGAAAAACA GGCCAAAGCC ATTAAGCAG CTGGCAGTGC
251 CAGGGGACAA TTGTGCCCA CGGTCTCAGC CTGGGCCTGT CACGAGCTTG
301 CAGAGTTAAG ACTCTGCCAC AGAGAAGAGA ACATCAGGAC ACCTGGCAGC
351 CCTATGCTTT ACAATGTGGC ATCCAGAACC CTCACCACC TCACTGTGCC
401 AGAGAAGTGG GCATGGCTGG GGTCCCCGTC GCCATTTGAC AGCAAAGACC
451 CAAGAGGATA GATGACACAC AGCATCTGGT GTCACACAGA CTGGGATTAG
501 AATCCAGGCA CGGTCCTTCA CTAGCTGTGT GACCTTGGGA AAAGGACTTG
551 ACTGTTCTGT GCCTCAGTTT CCCCATCTGT AAAACGGAGG CTAAAAAAT
601 ACTGATCGGA CACAGTGGTC AGGGTTAGAG ATAACATACA TGAAACGACC
651 ACAAGCTCCC CAAGGGCAAA GGTTCCTGAC ATCCCGGTC TCTGCCATTT
701 TCCATGTGCC CAGAAGAGCA CTTGGTCCAT AGTATGTGCT CAATGAATGT
751 AAATGGGATA AAAACACGAA CGAACACTCT GCCAACGATG CTGCTGTGCC
801 TTTGTCATCA CTGCTTCTGT TTAGGCTGTA GCTGACTTAT CTAAGGCCAT
851 ACAGCTGCTC AATGCATAGC CCGGCCAGTC CCG

C



gb43.30

gbDhDi43.30.seq (secuencia celular de gb43 más cerebro de ms adicional)

Longitud: 291

pcr larga según Temp.

cebador: t3pb 1+2 aso. 65%

vector : pCR2.1, Invitrogen

```

1  CCCCTTGACT TCGGTGTGTA AACTTGTGGT ATAGAACATG ATGTTTTAAG
51  ATACATGTAC ATTGTGGAAT GGCTTGATCA TGCTAATTAA CATATGAATT
101 ACCTCACTTA GCTATCTTTT TTATGGTGAA AGCACTTAAA ATCTACCCCTC
151 AGCAGTTTTC AAGTACACAA TACATTCTA TTAACATAG TCACCATGTT
201 GTACAATAAA TCTCTTGAAT TTATTCCTCC TGCCTAACTG ACATTTTGTA
251 TCCTTGACT GATCTCTCTC CCCAGTCCCG TGCCCGAATT G
    
```

ES 2 657 705 T3

Negrita - secuencia del cebador
Cursiva - continuación de la secuencia celular

BLASTN2 de: /home/vir088/ttbrain/gbDhDi43.30.seq desde: 1 a : 291 9 de Junio, 2010
12:21
comparada con la base de datos : **nrnuc**

>>>>nrnuc:GI_225543527 Gi|225543527|ref|NG_011635.1| **Miosina IIIA de Homo sapiens (MYO3A)**, RefSeqGene en el cromosoma 10. 0/0
Longitud = 285464

Puntuación = 525 bits (265), Esperada = e-146
Identities = 268/269 (99%)
Hebra = Plus / Plus

```
CCCTTGACTTCG
Query: 14  gtgtgtaaacttgtggtatagaacatgatgttttaagatacatgtacattgtggaatggc 73
          |||||
Sbjct: 276074  gtgtataaaactttgtggtatagaacatgatgttttaagatacatgtacattgtggaatggc 276133
          TAAATGACAAAAC

Query: 74      ttgatcatgctaattaacatatgaattacctcacttagctatctttttatggtgaaagc 133
          |||||
Sbjct: 276134  ttgatcatgctaattaacatatgaattacctcacttagctatctttttatggtgaaagc 276193

Query: 134     acttaaaatctaccctcagcagttttcaagtacacaatacatttctattaactatagtc 193
          |||||
Sbjct: 276194  acttaaaatctaccctcagcagttttcaagtacacaatacatttctattaactatagtc 276253

Query: 194     ccatgtgtacaataaatctcttgaatttattcctcctgcctaactgacattttgtatcc 253
          |||||
Sbjct: 276254  ccatgtgtacaataaatctcttgaatttattcctcctgcctaactgacattttgtatcc 276313

Query: 254     tttgactgatctctctccccagtcccgtg 282      CCCGAATTG
          |||||
Sbjct: 276314  tttgactgatctctctctccccagtcccgtg 276342  ACCAGTGCCCT
```

>>>>nrnuc:GI_14018255 Gi|14018255|emb|AL162503.12| Secuencia de ADN humano del clon RP11-420F12 sobre el cromosoma 10 Contiene el gen GAD2 para la glutamato descarboxilasa 2 (islotos pancreáticos y cerebro 65kDa), el extremo 3' del gen MYO3A para la miosina IIIA y dos islas CpG, . . . 0/0

Longitud = 175594

Puntuación = 525 bits (265), Esperada = e-146
Identities = 268/269 (99%)
Hebra = Plus / Plus

```
Query: 14      gtgtgtaaaactttgtggtatagaacatgatgttttaagatacatgtacattgtggaatggc 73
          |||||
Sbjct: 35112  gtgtataaaactttgtggtatagaacatgatgttttaagatacatgtacattgtggaatggc 35171

Query: 74      ttgatcatgctaattaacatatgaattacctcacttagctatctttttatggtgaaagc 133
          |||||
Sbjct: 35172  ttgatcatgctaattaacatatgaattacctcacttagctatctttttatggtgaaagc 35231
```

ES 2 657 705 T3

```

Query: 134  acttaaaatctaccctcagcagttttcaagtacacaatacatttctattaactatagtc 193
            |||
Sbjct: 35232 acttaaaatctaccctcagcagttttcaagtacacaatacatttctattaactatagtc 35291

Query: 194  ccatggttgacaataaatctcttgaattttattcctcctgcctaactgacattttgtatcc 253
            |||
Sbjct: 35292 ccatggttgacaataaatctcttgaattttattcctcctgcctaactgacattttgtatcc 35351

Query: 254  tttgactgatctctctccccagtcocctg 282
            |||
Sbjct: 35352 tttgactgatctctctccccagtcocctg 35380
    
```

Marco de lectura abierta : 49aa

(**Péptido**) **FASTA** de: gbDhDi43.30rev.49.pep desde: 1 a : 49 9 de Junio, 2010

REFORMAT de: gbDhDi43.30rev.49.pep check: 3689 desde: 1 a : 49 9 de Junio,

PARA: **SwissProtPlus**:* Secuencias : 11,223,768 Símbolos : 3,635,054,084 Tamaño de palabra: 2

```

SPTREMBL:Q9WB12_9VIRU  Begin: 19  End: 60
! Q9wb12 SubName: Full=ORF2; Flags: F... 84 109 108 181.5 0.3
SPTREMBL:Q9WB09_9VIRU  Begin: 19  End: 54
! Q9wb09 SubName: Full=ORF2; Flags: F... 91 104 176.0 0.61
SPTREMBL:Q9WB02_9VIRU  Begin: 19  End: 59
! Q9wb02 SubName: Full=ORF2; Flags: F... 82 82 100 168.8 1.5
SPTREMBL:Q9WSW4_9VIRU  Begin: 32  End: 73
! Q9wsw4 SubName: Full=ORF2; Flags: F... 82 82 97 162.5 3.4
SPTREMBL:Q9WAY4_9VIRU  Begin: 19  End: 54
! Q9way4 SubName: Full=ORF2; Flags: F... 84 84 94 160.8 4.2
SPTREMBL:Q9WB10_9VIRU  Begin: 19  End: 54
! Q9wb10 SubName: Full=ORF2; Flags: F... 84 84 94 160.1 4.6
SPTREMBL:Q9WAZ2_9VIRU  Begin: 19  End: 54
! Q9waz2 SubName: Full=ORF2; Flags: F... 84 84 94 160.1 4.6
SPTREMBL:O70807_9VIRU  Begin: 23  End: 58
! O70807 SubName: Full=Putative uncha... 84 84 94 158.2 5.9
\\End of List
    
```

gbDhDi43.30rev.49.pep
SPTREMBL:Q9WB12_9VIRU

```

ID  Q9WB12_9VIRU          Sin revisar;          150 AA.
AC  Q9WB12;
DT  01-NOV-1999,         integrado en          UniProtKB/TrEMBL.
DT  01-NOV-1999,         versión de secuencia 1.
DT  09-FEB-2010,        versión de entrada 20.
DE  SubName: Full=ORF2; . . .
    
```

```

PUNTUACIONES Init1: 84  Initn: 109  Opt: 108  puntuac.z : 181.5 E(): 0.3
>>SPTREMBL:Q9WB12_9VIRU          (150 aa)
    initn: 109 init1: 84  opt: 108  puntuac.z: 181.5 esperada(): 0.3
Puntuación de Smith-Waterman : 108; 53.5% de identidad en un solapamiento de 43 aa
(6-46:19-60)
    
```

ES 2 657 705 T3

```

                                10      20      30      40
gbDhDi43.30r                MFYTTSLHTEVKQFGHGTGERDQSKDKCQ--LGRRNKFKRFIVQH
                                ||||| ||||| ||| |::| | : ||: | || : |
Q9WB12_9VIRU AQTQRRVIPASRGRVPEVSLHTXVKGQFGLGTG-RAMGKALKKDMFLGKLYKKRALSLH
                                10      20      30      40      50

```

```

gbDhDi43.30r GDYS
                                |
Q9WB12_9VIRU GLRTPEAKPPAMSWRPPVHNPNRIERNLWEAFFRIHASSCGGHLVGHLLTVLARRYGAPP
                                60      70      80      90      100     110

```

```

ID   Q9WB12_9VIRU                Sin revisar ;           150 AA.
AC   Q9WB12;
DT   01-NOV-1999,               integrado en UniProtKB/TrEMBL.
DT   01-NOV-1999,               versión de secuencia 1.
DT   09-FEB-2010,              versión de entrada 20.
DE   SubName: Full=ORF2;
DE   Flags: Fragmento;
OS   Torque teno virus.
OC   Virus;                    virus de ADNss ; Anelloviridae; Anelloviridae sin clasificar.
OX   NCBI_TaxID=68887;
RN   [1]
RP   SECUENCIA DE NUCLEÓTIDOS.
RC   TEJIDO=Sero;
RX   MEDLINE=99335592; PubMed=10405352; DOI=10.1006/viro.1999.9797;
RA   Hijikata M., Takahashi K., Mishiro S.;
RT   "Complete circular DNA genome of a TT virus variant (isolate name
RT   SANBAN) and 44 partial ORF2 sequences implicating a great degree of
RT   diversity beyond genotypes.";
RL   Virology 260:17-22(1999).
CC   -----
CC   Derechos de autor de UniProt Consortium, véase http://www.uniprot.org/terms
CC   Distribuido bajo la licencia de Creative Commons Attribution-NoDerivs
CC   -----
DR   EMBL; AB024379; BAA77446.1; -; Genomic DNA.
DR   InterPro; IPR004118; Gyrovir_VP2/TT_ORF2.
DR   Pfam; PF02957; TT_ORF2; 1.
PE   4: Previsto;
FT   NON TER                    1          1
FT   NON TER                    150        150
SQ   SECUENCIA 150 AA; 16415 MW; A4A7DF1585FC40D CRC64;
    AQTQRRVIPA SRGRVPEVSL HTXVKGQFGL GTGRAMGKAL KKDMFLGKLY KKRALSLHG
    LRTPEAKPPA MSWRPPVHNP NRIERNLWEA FFRIHASSCG CGHLVGHLLTV LARRYGAPP
    PPAPGAPRPA LKRQLALPAP PADPQQANPT

```

Figura 9

A

hod11

hodL.VvWw.1.seq Longitud: 639
Hodgkin L1236+TPA

Cebador de PCR: h1f1+Lr1 GCI 59°C

Vector pGEM T-Easy

```

1  CCCCTTGACT GCGGTGTGTA AAGCGCCCCA GCCTGTGCCT GCACAGTGCC
51  TGTGTGGTGT GAACCCATGA CCAGGCCTCT GGAGGGAAGG AAGGTTAGGC
101 TTAGTGGACA CCAGCTTTC TAAGGTGGGT CTTAGACCAA CTCATTAAAA
151 TGGCAGGATG GCCTTTTGTG CTGTATTCT TGGGATTTTC AAGATGCCCC
201 ACACAGCAGA AGGGATGTGC ATTTTTTCT CTGCCCTGAG TTGTTTGATA
251 AAAATCAGTG ACCTCGTTCT CCACTTAGAA CTCCCCTGAA CTGCACTCGG
301 TGTCTAGGAC TGTGGGGAA GGAAGTGAAG AGCCAGCATG TAGTCTCCTC
351 TGGACTCTTA CAGGATCTGT CCACCTCTGG GCTCTTTATG TAGGGGAAGG
401 TGTGAGCTCC TGGGAGTACT CCTGATAGAG GACTGTTTCC CTGAAAACCT
451 CAGCAGTGTG TGAGGCCCTA GCAGGGGGAA CCCAGACCCC GCCTGCCAAA
501 GCCCCTAATC CCTCAGGGCT ATTATCAGCA GCCTAAGCGC CTTAGGGTGG
551 CCAGAGTCCA GCCCAGCAAG CAGCAAAGTC AGCAGCCTCC TCGCCCTATC
601 CTCTCCATGC CCCGGGGCAC TCCAGTCCCG ACCGAATTG
    
```

Cebadores en negrita

Continuación de la secuencia celular en cursiva

BLASTN2 de: /home/vir088/ttmixture/hodL.VvWw.1.seq desde: 1 a: 639
comparada con la base de datos: nrnuc ..

Base de datos: nrnuc

```

>>>nrnuc:GI_18121492 Gi|18121492|emb|AL513485.10| Human DNA sequ... 690 0.0
>>>nrnuc:GI_149944898 Gi|149944898|gb|AC198797.3| MACACA MULATTA... 509 e-141
>>>nrnuc:GI_55416061 Gi|55416061|gb|AC121551.11| Mus musculus ch... 84 1e-12
>>>nrnuc:GI_241752248 Gi|241752248|ref|XM_002400984.1| Ixodes sc... 46 0.26
>>>nrnuc:GI_291191454 Gi|291191454|gb|GU722348.1| Torque teno vi... 44 1.0
>>>nrnuc:GI_291191453 Gi|291191453|gb|GU722347.1| Torque teno vi... 44 1.0
>>>nrnuc:GI_291191452 Gi|291191452|gb|GU722346.1| Torque teno vi... 44 1.0
>>>nrnuc:GI_217416834 Gi|217416834|gb|FJ426280.1| Torque teno vi... 44 1.0
    
```

```

>>>nrnuc:GI_18121492 Gi|18121492|emb|AL513485.10| Secuencia de ADN humano
del clon RP11-48O20 sobre el cromosoma 1 Contiene el gen TAGLN2 para la
transgelina 2, el gen IGSF9 para el miembro 9 de la superfamilia de inmuno-
globulinas, el gen SLAMF9 para el miembro 9 de la familia SLAM, un gen
novedoso y . . . 0/0
Longitud = 80398
    
```

ES 2 657 705 T3

Puntuación = 690 bits (348), Esperada = 0.0
Identities = 348/348 (100%)
Hebra = Plus / Plus

Query: 282 tcccctgaactgcactcgggtgtctaggactggtgggaaggaagtgaagagccagcatgt 341
|||||
Sbjct: 26212 tcccctgaactgcactcgggtgtctaggactggtgggaaggaagtgaagagccagcatgt 26271

Query: 342 agtctcctctggactcttacaggatctgtccacctctgggctctttatgtaggggaaggt 401
|||||
Sbjct: 26272 agtctcctctggactcttacaggatctgtccacctctgggctctttatgtaggggaaggt 26331

Query: 402 gtgagctcctgggagtactcctgatagaggactgttccctgaaaacctcagcagtggtt 461
|||||
Sbjct: 26332 gtgagctcctgggagtactcctgatagaggactgttccctgaaaacctcagcagtggtt 26391

Query: 462 gaggccctagcagggggaacccagacccccctgccaaagcccctaatacctcagggcta 521
|||||
Sbjct: 26392 gaggccctagcagggggaacccagacccccctgccaaagcccctaatacctcagggcta 26451

Query: 522 ttatcagcagcctaagcgccttagggtggccagagtcagcccagcaagcagcaagtca 581
|||||
Sbjct: 26452 ttatcagcagcctaagcgccttagggtggccagagtcagcccagcaagcagcaagtca 26511

Query: 582 gcagcctcctcgccctatcctctccatgccccggggcaactccagtccc 629 **GACCGAATTG**
|||||
Sbjct: 26512 gcagcctcctcgccctatcctctccatgccccggggcaactccagtccc 26559 **AGCTGGCTGATC**

Puntuación = 505 bits (255), Esperada = e-140
Identities = 272/280 (97%)
Hebra = Plus / Minus

CCCCTTGA
Query: 9 **ctgcggtgtgtaaa**gcgccccagcctgtgcctgcacagtgcctgtgtggtgtgaacccat 68
|||||
Sbjct: 29128 **ctgctgtgtgtaaa**gcgccccagcctgtgcctgcacagtgcctgtgtggtgtgaacccat 29069
TTCAGTTAG

Query: 69 gaccaggcctctggaggaaggaaggttaggcttagtggacaccagctttcctaaggtgg 128
|||||
Sbjct: 29068 gaccaggcctctggaggaaggaaggttaggcttagtggacaccagctttcctaaggtgg 29009

Query: 129 gtcttagaccaactcattaaaatggcaggatgggcttttgtgctgtatcttctgggattt 188
|||||
Sbjct: 29008 gtcttagaccaactcattaaaatggcaggatgggcttttgtgctgtatcttctgggattt 28949

Query: 189 tcaagatgccccacacagcagaaggatgtgcannnnnnctctgccctgagttgtttga 248
|||||
Sbjct: 28948 tcaagatgccccacacagcagaaggatgtgcatttttctctgccctgagttgtttga 28889

Query: 249 taaaaatcagtgacctcgttctccacttagaactcccctg 288
|||||
Sbjct: 28888 taaaaatcagtgacctcgttctccacttagaactcccctg 28849

B

hoht33

hoHT.vAf.33a.seq Longitud: 3387

Linfoma HSB+TPA, Templiphi-RCA

Cebador de PCR HLF1+Hr4 60°C

Vector pCR2.1

```

1  AATTCGGTCG GGACTGGCAG AGTGACGCTC AGGTCAGCCT GACAGCAGGG
51  TGATTGAAGG GGCCAGATAC CCCAGCAGGG CCTGAGGCCA GAACACAGCA
101 TAGGCTGGCT CTGATGGGTG GAGGAGGTGG CCAGGCATCA TCTGGAGCTT
151 GGAGTTGAGA ACATCTGTGA CTCCTCCTTC AGGAGGGTGC TCTAGGAGTT
201 GAGAGCATCC TAGGTAGGAC CATACATCTA CCCCATCCT AGTCCCTCC
251 AGCCTCTCTT TTCAGCTCCA GGTCTACCTT AAGGGACCTA GGACACCTGG
301 GCTGGGGCAT AACAGGACTT GGTTTTATGT AAAGGAGCTG GGAAGAGACT
351 GAGATAACAG AGGGCTGCAA GGAGAGAGAC AGAGAGAGAA GAACCTGCCA
401 GAAGAAGCTC CTCAGCAATC CACTAAGCCC TGATCTTTGC CTCACTGCCT
451 GTCCTTCCC ATCCGCTCTT CTGCTCTCTC AATCTCTGCC TTCAAGAAAT
501 TTGGTGATA TTGGAATAGG GAGGAATAGA AGCACCCCTGG GTGGAGCTCT
551 GGGCTTGGCT GTGCACGAGC TTTCAGTGGG TGGTTGCTG GTCTCCAAAG
601 ATGACCCTCC ATTAGTCATG CTTCTCGGTG TTTGTCCTCA GGTAGTCTCA
651 TCCCATCTTG AGTCTGGGCT TGCCCTGTGA CTCACTTTAA CCACAAGAAT
701 GTGGCAGAAA GGATGTTGTG CCAGTTCTAG AACTAAGCCT TCAGAAAGCC
751 TAGCACCTTC TGCTTTTAGG AGCACTGAGC CCCCATGTTA GAAGTCCACT
801 TTTATACTCT GCTCTGGAGA CTAGCAGAAAT TAGAAATGCA CTGCTGAATG
851 CTGCTCGAGA GACTAATGGA GAGGCCATGT GAATAAGGAG GCCTGAAACT
901 ACATGGAGAT AGAGGGCCAG CCACCCAGC ACCACGGCTC AGCTGTGCCT
951 CCCAGCCATC TCTGCCAGTC CTCAGGGCT ATGAGTGAAC CATCTGGAT
1001 GTTCTAGCTC GGTGGAGCCC CCAGGTGATT GCAGCCTCAG CCACCATCTG
1051 ACTGTAGCTG CATGAGAGGC CCCCAGTGG ACCAGCAGGA CTGCCAAGCT
1101 GAGCCCTGCC CACCCACAGA ACTGTGAGAA ATAAAAAAT GGTGTTTCC
1151 TTAAGCCATT AAGTTTGGG ATGATTTGTT ACTACAATT GATAACTGAT
1201 ACAGTCTGTC TTTAGGAAA ACAAGGGATA ACTCTGGGCT CCAGGTGTCT
1251 TCTATAGGAT GAATGGGACT TGGTTGCTGA CAAGCTGACA AGTTTGAGCA
1301 TGAAACTCTT TTTTTTTTTT GGAGAAGGAA TTTTGCTCTT GTTATCCAGG
1351 CTGGAATACA GTGGTGGGAT CTCGGCCCAA GGCAACCTCT GCCTCCTGGG
1401 TTCAAGCAAT TCTCTGCCT CAGCCTCCTG AGTAGCTGGG ATTACAGGCA
1451 CCCACCACTA CACCTGGCTC TTTTTTTTTT TTGFATTTT AGTAGAGACA
1501 GGTTTTCATT ATGTTGGCCT GGTGAGGTTT TGAACCTCTG ACCTCAGGTG
1551 ATCCACCTGC CTGGCCTCC TAAAATGCTG GGATTACAGG TGTGAGCCAC
    
```

1601 CGTGCCTGGC CTGAGCATGA AACTTTTATG CTCAAACATT AAAGTGAAAA
 1651 CACTCACCAG CTCAGTGAA TAAGAACTTC TGGGGGCAAG GCCCAGGAAT
 1701 CTACAGTTTA GTAAGTGCCC CCACCACTGG ACCCTGGGAA AGTGGACTGC
 1751 ATTTGAAAA ACTCTAGATC AGTTGATACC CAGGAGTCCT CATAACACTA
 1801 AGTTGTAATA CCTCAGTGTG AATTAGTCTG ATGCAGCTCT TCTTAGAGGT
 1851 CATTGACAGA GGGCAAGACA TTCCAAAAG GAAGGAATAG CCAATATGGA
 1901 ATGACAGGTG GATTGGATGA CCCCTATTA TTTAGTTTCA ACCTGCCCTT
 1951 CTTGCCTTCC CTCCCACAAA TTCCCTTCA GATCCTCCGT CCTAATCCTC
 2001 TTCGATAGTT CAITGTTCTT CTGCAGACAG AGCAGCGAAG TGTATCTGT
 2051 TGTACCCACT ATGACTAGTT GATGGTGCAT GGCTTCCATG GAGCACTGCT
 2101 GTGATCCATT AGTCATGGAG CAGTGTCTG ATCCATTGTC ATGTCTGCCA
 2151 TGAACACTGG AAGGGGAGT GGTAAATGACA GCCTCTTACA TTTGCCAACT
 2201 CTGCCCAACA TTCTTCCCAG TGTGGGAAA GCCTTTGCTT ATTCCATCC
 2251 TTCTTGGAAA GCTTTCTCC FCCATTTCC ATTTTAAAT TTTCTCATT
 2301 TTATGGTGA CCATGGATAC CACCTGTCCA TATAGCTGGC TTCTGATTT
 2351 TCCAGATGAA AGTAATCCTT CCTCTCCTAA CCTCCATGA CACCTAACCT
 2401 GGCACCTATT TACGGTGTTC AGCTCCTTCT CTTGACGTT CTCATTGTTT
 2451 TCCTCTCATC TTCTCCCAG GAATGGATC CCCGCCAAGG GAGGTACCAG
 2501 GTCAGTTTCT TCTTTGTGCA ACAGGGTGTG CCTGATGAGC ACAAACCTGG
 2551 AACAAAGTGT TGTAGGGCTG GTGGGCATCT GGTCTCTCTG GGTGTGTGT
 2601 AGCCTGAGCC GGGGGGCAA TGGGTGTTG TTTTCTGAA GAAGGCAGGC
 2651 GTCTGTGGC AGATGIGGGT GGAGGGGTT GGGAGTAGT ATCATGGAGA
 2701 GGCTGGGATC CTATCTATCT CCTTCCCCTG CTTGAAGGGC AACTTGGGAG
 2751 AAGCTCAAGA GGGAGGAGT GACTGCAGAA GCTGGGATAC CTGCATAACT
 2801 CTCAGGTCA AGCATCACTG CTTAGGGCC CTGGGGCCT ATGTGTGAGT
 2851 CAAGAAAGGG AGATAGAGAG AGAAGAGAGA GAGAGGAGAG AGAGAGAGAG
 2901 AGAGACAAGA CAGAGGAGAG AGAGAGAGAA GAAAGAGGAG AGAGAGAGAA
 2951 GAGAGAGCAG AGAGAGAGAG CATGCTGTCA GTGAGGTGGC CCTAAGCCCT
 3001 CTTGAAATA ACTTGGAGGC ACTGTGGGT GGCTCTGAGG TGCTGAGGTA
 3051 TACCTGTAGT GGGCTIAGGA CCTTTCCAAC CTGGGTCTGA AGGTTGAGGC
 3101 AACCTGGGT GTACCTGCTG GTGAGCTGAG AGCCCTGGGG ACCTTGGCA
 3151 GACATTCCCA CCCCTGCAGC CTGGAGGTT TGATGCAGT GAGGCTGTCC
 3201 TGCTCATCAC TACGTCTCTT GGGACAGCAC ATTGCCTGTG CTGAACAGGC
 3251 ATTCAGTTGC GATTTGTGGA ATCAGTGTG GTGAGGAGGG CAAGTGGCAA
 3301 CAGAAATGGG GGTGTGCTCC CCCAGTTCC TCAGCTACAA TCTCCATGAC
 3351 CTTCTACTACT GCCCTGGGCC CAGTCCCAGC CGAATTG

C

ES 2 657 705 T3

hoht22
hoHT.vAF.22a.seq Longitud: 1790
Linfoma HSB+TPA, Templphi-RCA
Cebador de PCR HLF1+Hr4 60°C
Vector pCR2.1

```
1 CAATTCGGTC GGGACTGGG AGCTGTGAGA AAGAGAAGAG AAGTCAGAT
51 CAGGAACATT ACACAGAAGT CGGCAAACT GGAACGAGGA GGGAAAGAAA
101 TGAGCGAGTC TGACACTCAG TCCATCCTAG TTCCTATCAC ACAGGGAGGG
151 ACATTGCCAT GCACATCCCC ACAGAGATGC ACCGTGTAAG GGGTCGAGGC
201 AGATCCTGTC CACTATTGCC AGCTCTGAGG TGATCAAATT GTGCTCGCCC
251 AGGGTAACCC GGTGACCTA AACCAACCCA CTCCTTGCA CATCTTAGGT
301 GTTCCTGAGT CAGCAAGGCT GAGGAAGCCA CTCCAGCCAA AATCCCTTGT
351 GCGATCTTCA AGCCCCAATC ACAGGCAATG ACAAGGCCAT GTCTGGCTGG
401 CCTCATGGGG ACTGCCCTCC CCTCACCAGA CCTAGAACAC AGGCAATGCT
451 CAGCAGCGTT CTGAGAAGAG CTGAGGTCAA GAACTCCAAC CCCACGCAAC
501 CCAGACCTGA TACAACAGA CACCCATTG CACTCCTAAC CCTTGAGCCT
551 CTATTTCCAG ACCTCCTCAC TGGGTCTCAG CTGAGAACCC ACTTTTAGCC
601 AAGCATCTT AGTTCAGAGT TCCTCGCAGT GAGGGGATCC CTCCTTGCC
651 TTGCTGTCTG TGCTGCATCC ATTATACCCT CACACCGTGC TACTCAGCAG
701 GGGAGAAATG GAGCCCTGGG GAGCCGGCAC TTTTCTCTTC TGCCTCTTCC
751 TTGCCTTGCC TCAGGAAGGG GAAAACTCT GGGTTGTTTT AGTTTGATCC
801 CCTGTCTTAA GTGACCACAG GAACACTAGG CAGTGAGTAC ATATGGATTC
851 TTAGCAGAGA GCTGACAAGT CTTCAGAAAC ATAGAAAACA TAGAAGCTTT
901 GAGTGAGGAG ATCAGAATGT AATTAGGAGT TTCTTTTGA GCAAACCCCA
951 CCCCAGAGA GTGAGCCAA GTTCTTGAAG GCCCACCTGA GCAGATGACA
1001 CCAGCGTCTT CACTATGGCC ACAGTTGTGG GTGAGCCAGC CATTGTGGGG
1051 GCAGCTCCAC AGGTAGGACT CGTGTCTGA GCAGGCACA TCATCCAGGA
1101 CAATGGGTCC TGAGCCCTGG CCAACTGGG CATTTCCTGG GGCTGACATG
1151 GCCCAGCCAC AGCCCGGCTG CCTGCAGACC ACATTGGCAT CATTGGTGTC
1201 CCAGTAGTCA TCACACACGG TGCCCCAGGA GCCTCGGTAT AGGACCTCCA
1251 CTCGGCCTCG ACACCTGTCG CCTCCATTCA CCAGCCTCAG GGCCAAACTG
1301 GATTGAGATC CTACAGGGGA ACACAAGAAC CTTTCATCCA TCCCTATCAT
1351 GAGGTCAAGA ATCTAAGTA AGTTCCACAC TCAGGGTACT TCCTAATGAA
1401 CTAAGTCACC TAGGCAGGCA GTCACCTTG CATATGACTA CAGACTAGGC
1451 TTCATACCG TGAAGTAGC ACTGATAACC TACTCTGCC AGGTCTATGG
1501 GTGCTCAACT TTTGGGGAAG CACCTGTGAC CCCAGTGGAT GTGATGGGAA
1551 TGGATGCCCC ACTCCCAGT TGGGTACACA GAGGATGGAG CTGCTCAGCT
1601 CCAGATGGCA GGCCAGACC CCTCCCTTAT TCAGGAGCAT GGTCTATCT
1651 GGGATCTGAC TGGCAGAGTA CCAGAGATGG CAGGGATGAG GTCCCCATAG
1701 GATTAGGGAG ACCCCAGGG CTGTCTGTA GCCCATAGAT AAGGATCTTT
```

1751 TCTGACCACT TGGAACAGGA TCCCAGTCCC GACCGAATTG

Figura 10

DhDi:

t3pb-1.prime Longitud: 18 cebador directo de pcr larga
1 caattcgggc acgggact

t3pb-2.prime Longitud: 24 cebador inverso de pcr larga
1 ccccttgact tcggtgtgta aact

cd:

tth4prime1.seq Longitud: 28 cebador directo de pcr larga consecutivo con prime2
1 CAGCGAGAAC GCCACGGAGG GAGATCCT

tth4prime2.seq Longitud: 28 cebador inverso de pcr larga consecutivo con prime1
1 CGGACGGGCG TGGAAACTC AGCCATTC

DfDg:

hel32-1.prime Longitud: 18 cebador directo de pcr larga
1 cgggactggc cgggctat

hel32-2.prime Longitud: 19 cebador inverso de pcr larga
1 agcccgaatt gcccttga

Figura 11

A

ttgb33.35

gbDhDi33.35.seq (genoma completo de una biopsia de ms)

pcr larga , Templiphi

cebador: T3PB 1 + T3PB 2 aso.:65 C

vector:TA-cloning pCR2.1 invitrogen

Longitud: 3725

```

1  ATTTTGTGCA GCCCGCCAAT TTCTGTTCAA ACAGACCAAT CAGGACCTTC
51  TACGTGCACT TCCTGGGGCG TGTCTACGAG GTCATATAA GCAACAGCGG
101 TGACGAATGG TAGAGTTTTT CTTCGCCCGT CGCGGGCGAG AGCGCGAGCG
151 AAGCGAGCGA TCGAGCGTCC CGTGGGCGGG TGCCTAGGT GAGTTTACAC
201 ACCGAAGTCA AGGGGCAATT CGGGCACGGG ACTGGCCGGG CTATGGGCAA
251 GGCTCTTAAA AAATTCCCCC GCTCTGCTCT CCGGCAGGAC ACAAAGTCAT
301 GCCGTGGAGA CCGCCGGTCC ATAACGTGCC AGGTAGAGAG AATCAATGGT
351 TTGCAGCGTT CTTTCACGGT CATGCTGCTT TCTGCGGGTG TGGTGACCCT
401 GTTGGGCATC TTAACGGCAT TGCTCCTCGC TTTCCTAACG CCGTCCACC
451 GAGACCACCT CCAGGGCTAG ACCAGTTAA TCCCAGGGC CCGCAGGTC
501 CCGGAGGGCC CCCCGCCATC TTGCCAGCTC TGCCGGCCCC GGCAGACCCT
551 GAACCGGCAC CAGGCGTGG TGGTGGGGCA GATGGAGGGC CCGCCGTGG
601 GGCCGCCGCC GACGCAGACC ATACCGGTA CGAAGAAGGA GACCTAGAAG
651 ATCTTTTCGC CGCCGCGGCC GAGGACGATA TGTAGTAGG CCGAGGCGCC
701 GCCGCTACTA CAGGCGCAGA CTGAGACGGG GCAGACGAG AGGCGACGA
751 AAGAGACACA GACAGACTCT AGTAGTGAGG CAGTGGCAAC CTGACGTTGT
801 TAAAAAGTGT AAAATAACAG GATGGATGCC TCTTATAATC TGTGGCTCTG
851 GAAGCACACA GATGAACTTT ATAACTACA TGGACGATAC TCCCCCTATG
901 GGATACACCT ACGGGGGCAA CTTGTAAAT GTAACCTTCA GTCTAGAGGC
951 CATCTATGAA CAATTCCTGT ACCACAGAAA CAGGTGGTCC AGGTCTAACC
1001 ATGACTTAGA CCTGGCCAGA TACCAAGGAA CCACTCTAAA ACTTTACAGA
1051 CACCAAACCG TGGACTATAT AGTTAGTAC AACAGAACAG GCCCCTTTAC
1101 TATAAGTGAA ATGACTTACA TGAGCACACA CCCGGCTCTC ATGCTACTAC
1151 AAAAACATAG AATAGTTGTA CCCAGCTTCA GAACCAAGCC AAAAGGCAAA
1201 AGAGCCATAA AAATTAGAAT AAGGGCCCCA AACTAATGC TCACCAAGTG
1251 GTACTTTACA AAAGACATT GCTCCATGGG CCTTTTCAA CTAATGGCAA
1301 CAGCTGCAGA ACTTACAAAC CCATGGCTCA GAGACACCAC AAAAAGCCCA
1351 GTAATTGGCT TCAGAGTCTT AAAAAACAGC TTATACACAT GCCTTTCCAA
1401 CTTAAAAGAC CAAGCAATAC AAGGTGAAAG AAAGACTGTA CAAAATAGAT
    
```

ES 2 657 705 T3

```

1451 TACACCCAGA AAACCTACAT GGCACAGGAC CTAATGCTAA AGGCTGGGAA
1501 TACACATACA CAAACTAAT GGCATCTACA TACTACTCAG CCAACAGAAA
1551 CAGCACCTAC AACTGGCAAA ACTATCAAAC TAACCTATGCA AACACATATA
1601 CAAAATTTAA AGAAAAAAGA ACAGCAAAC TAACTTAAT TAAAGCAGAA
1651 TACCTATATC ATTACCCTAA CAATGTCACA CAATCTGACT TTATATTAGA
1701 CTACACACTA ACACCCGACT GGGCATATA CAGCCCTAC TACCTAACAC
1751 CCACCAGAAT TAGCCTAGAC TGGGACACAC CATGGACATA TGTAAGATAC
1801 AACCCACTAT CAGACAAAGG CATAGGTAAC AGAATATATG CACAGTGGTG
1851 CTCAGAAAAA TCTAGTAAAT TAGACACCAC AAAGAGCAAG TGCATACTAA
1901 GAGACTTCCC ACTGTGGGCC ATGGCCTATG GCTACTGTGA CTGGGTGGTG
1951 AAGTGCACAG GAGTGTCCAG TGCTTGGACA GACATGAGAA TAGCCATTAT
2001 ATGTCCCTAC ACAGAACCAG CACTTATAGG GTCACAGAA GACGTAGGCT
2051 TCATCCAGT AAGTGACACC TTTTGCAACG GAGACATGCC GTTCTTGCA
2101 CCATACATAC CTATTACATG GTGGATTAAG TGGTACCCCA TGATTACACA
2151 CCAAAGGAA GTTCTTGAGG CAATAGTTAA CTGTGGACCG TTTGTACCCC
2201 GAGACCAAAC TTCCTCAGCT TGGGAATAAC CATGGGTAC AAAATGGATT
2251 GGAANTGGGG CGGCTCTCCC CTGCCCTCAC AGGCAATCGA CGACCCCTGC
2301 CAGAAGTCCA CCCACGAACT TCCCGACCCC GATAGACACC CTCGCATGTT
2351 ACAAGTCTCT GACCCGACAA AGCTCGGACC GAAGACAGT TTTCACAAAAT
2401 GGGACTGGAG ACGTGGGATG CTTAGCAAAA GAAGTATTAA AAGACTCCAA
2451 GAGACTCAA CAGACGATGA ATATGTTGCA GGACCCTTAC CAGAAAAAG
2501 AAACAAGTTC GATACTCGAG TCCAAGGCC TCCAACCCCA GAAAAAGAAA
2551 GTTACTTTT ACTCCAAGCC CTCCAAGAGT CGGGCAAGA GAGCAGCTCA
2601 GAGGACCAAG AACAAAGCACC CCAAGAAAAA GAGGACCAGA AGGAAGCGCT
2651 CATGGAGCAG CTCAGCTCC AGAAACACCA CCAGCGAGTC CTCAGCGAG
2701 GCCTCAAACCT CCTCCTCGGA GACGTGCTCC GACTCCGGAG AGGAGTCCAC
2751 TGGGACCCCC TCCTGTCTTA ATTCAAGGTC CCAGTATCCC AGACCTGCTT
2801 TTCCCTAACA CACAAAAAA AAAACGATTT TCCAACACG ACTGGGTGTG
2851 CGAGTACGAG CTGGCCAAAT GGATGGATCG GCCCTTGGCG CACTACCCAT
2901 CAGACCCCCC TCACTACCCC TGGCTACCAA AAAAGCCTCC TACCCCTCCT
2951 ACATGTAGAG TAAGTTTCAA ATTAAGCTC AATGACTAAA ATTCAAGGCC
3001 GTGGGTGTTT CACTTCATCG GTGCTACCT CTAAAAGTCA CTAAGCACTC
3051 CGAGCGTAAG CGAGGAGTGC GACCCCCCTG CCCGGTAGCA ACTTCTCGG
3101 GGTCCGGCGC TACGCCTTCG GCTGCGCGG GCGCCTCGGA CCCCCCTCG
3151 ACCCGAATCG CTCGCGGAT TCGGACCTGC GCCTCGGGG GGGTCGGGG
3201 CTTTACTAAA CAGACTCTGA GGTGCCGTTG GACTGAGG GGGTGAACAG
3251 CARGAAAGT GAGTGGGGCC AACTTCGCC ATAAGGCCTT TAACTTGGG
3301 TCGCTTGTC A GCAGCTTCCG GGTCCGCTG GAGGCCGCA TTTTACATTC
3351 GGCCGCCATT TTAGGCCCTC GCGGGCCTCC ATAGTCGCAC ATCAGTGACG

```

ES 2 657 705 T3

```
3401 TCACGGCAGC CATCTTGGCT GTGACGTCAA CGTCACGTGG GGAGGACGGC
3451 GTGTAACCCG GAAGTCATCC TCATCACGCG ACCTGACGTC ACGGCCGCCA
3501 TTTTGTGCTG TCGCCATCT TGTGACTTCC TTCCGCTTTT TGTAATAAAA
3551 AGAGGAAGTG TGACGTAGCG GCGGGGGGGn nnnnnnnnnn nnnnnnnCGC
3601 CACCAGGGGG CGCTACGCGC CCCCCCGCG GCATGTGCGG GTCCCCCCCC
3651 TCGGGGGGGG CTCCGCCCCC CCGGCCCCCC CCGGGGCTAA ATACACCGCG
3701 CATGCGGGGC CACGCCCCCG CCGCC
```

B

zpr4.20
zpr4.20.seq (molécula subviral , ttgb33.35)
293TT +4 (gbDhDi33.35) PCR
Tr9.7 A/9nested B5

en pCR2.1, Nova Blue

zpr4.B5.20.seq Longitud: 719

```
1 CAATTCGGGC ACGGGACTGG CCGGGCTATG GGCAAGGCTC TTAAAAAATT
51 CCCCCGCTCT GCTCTCCGGC AGGACACAAA GTCATGCCGT GGAGACCGCC
101 GGTCATAAC GTGCCAGGTA GAGAGAATCA ATGGTTTGCA GCGTTCTTTC
151 ACGGTCATGC TGCTTTCTGC GGGTGTGGTG ACCCTGTTGG GCATCTTAAC
201 GGCATTGCTC CTCGCTTTC TAACGCCGGT CCACCGAGAC CACCTCCAGG
251 GCTAGACCAG CTTAATCCCG AGGGCCCGGC AGGTCCCGGA GGGCCCCCGG
301 CCATCTTGCC AGCTCTGCCG GCCCCGGCAG ACCCTGAACC GGCACCACGG
351 CGTGGTGGTG GGGCAGATGG AGGCGCCGCC GCTGGGGCCG CCGCCGACGC
401 AGACCATAAC GGTACGAAG AAGGAGACCT CGGGGGGGGC TCCGCCCCCC
451 CCGCCCCCCC CCGGGCTAAA TACACCGCGC ATGCGCGGCC ACGCCCCCGC
501 CGCCATTTTG TGCAGCCCGC CAATTTCTGT TCAAACAGAC CAATCAGGAC
551 CTTCTACGTG CACTTCCTGG GCGTGTCTA CGAGGTCTAT ATAAGCAACA
601 GCGGTGACGA ATGGTAGAGT TTTTCTTCGC CCGTCCCGG CGAGAGCGCG
651 AGCGAAGCGA GCGATCGAGC GTCCCGTGGG CGGGTGCCGT AGGTGAGTTT
701 ACACACCGAA GTCAAGGGG
```

Figura 12

A

tth25 (genoma completo)
Longitud 3758

```

1  AAGTACGTCA CTAACCACGT GACTCCCGCA GGCCAACCAG AGTCTACGTC
51  GTGCACTTCC TGGGCATGGT CTACATCATA ATATAARGAAC GTGCACTTCC
101 GAATGGCTGA GTTTTCCACG CCCGTCCGCA GCGAGAACGC CACGGAGGGA
151 GATCCTCGCG TCCCGAGGGC GGGTGCCGGA GGTGAGTTTA CACACCCGAG
201 TCAAGGGGCA ATTCCGGGTC GGGACTGGCC GGGCCCCGGG CAAGGCTTCT
251 AAAAAATGCG TTTTCGACAG GTTGCCGAGA AAAGGAAAGT GCTTTTGCAA
301 ACTGTGCCAG CTGCAAAGAA GGCTAGGCGG CTCTAGGTA TGTGGCAGCC
351 CCCCACGCAC AATGTCCCGG GCATCGAGAG AAAC TGGTAC GAGAGCTGTT
401 TTAGATCCCA CGCTGCTGTT TGTGGCTGIG GCGATTTTGT TGGCCATCTT
451 AATCATCTGG CAACTACTCT GGTGCTCCT CGCGTCCCTG GGGCCCCAGG
501 CGGACCCCGC ACGCCGAAA TAAGAAACCT GCCAGCGCTC CCGGCGCCCC
551 AGGGCGAGCC CGGTGACAGA GCGCCATGGC ATGGGGCTTC TGGGGCCGAC
601 GCCGCCGGTG GAGACGATGG AGAGCGCGGC GCAGACGGTG GAGACCCGCG
651 AGACGTAGGA GACGACGCC TACTCGCCG TTTGAGCTC GTCGAAGAGT
701 AAGGAGCGCG GGGGGGAGGT GGCGCAGACG CTACAGAAAA TGGCGACGGG
751 GCAGACGCAG ACGGACTCAT AGAAAAAGA TAGTCATAAA ACAGTGGCAA
801 CCAAACCTTA TAAGACGCTG CTACGTCATA GGGTACTTAC CACTTATATT
851 CTGCGGGGAA AATACAACCG CCCAGAACTT TGCCACTCAC TCGGACGACA
901 TGATAAGCAA AGGACCGTAC GGGGGGGGCA TGAATCCAC CAAATTCAC
951 CTGAGAATAC TGTACGACGA GTTACCAGG TTTATGAACT TTTGGACTGT
1001 CAGTAACGAA GACCTAGACC TGTGTAGATA CGTGGGCTGC AAAC TAATAT
1051 TTTTAAACA CCCCACGGTG GACTTTATAG TACAGATAAA CACTCAGCCT
1101 CCTTCTTAG ACACGCACCT CACCGCGGCC AGCATAACCC CGGGCATCAT
1151 GATGCTCAGC AAGAGACACA TACTAATACC CTCTCTAAAG ACCCGGCCCA
1201 GCAGAAAAACA CAGGGTGGTC GTCAGGGTGG GCGCCCCAAG ACTTTTTCAG
1251 GACAAGTGGT ACCCCAGTC AGACCTGTGT GACACAGTTC TGCTTTCCAT
1301 ATTTGCAACC GCCTGCGACT TGCAATATCC GTTCGGCTCA CCACTAAGT
1351 ACAACCCCTG CGTCAACTTC CAGATCCTGG GGGCCAGTA CAAAAACAC
1401 CTTAGTATTA GCTCCACTAT GGATCAAACT AACGAAAACC ATTATAARGA
1451 AAAC TTATTT AACAAAACTG AACTATACAA CACCTTTCAA ACCATAGCTC
1501 AGCTTAAAGA GACAGGACAC ATTTCAGGCA TTAGTCCTAC TTGGAATGAA
1551 GTCCAGAATT CAACAACACT TACTAAAGGA GGTGACAATG CCACTCAGAG
1601 TAGAGCACT TGGTATAAG GAAATACATA CAACGAGAAG ATATGCGAGT

```


ES 2 657 705 T3

1651 TAGCACAAAT AACCAGAAAC AGATTTAAAA ATGCAACCAA AGGAGCACTA
 1701 CCAAACTACC CCACAATAAT CTCACAGAC CTATATGAAT ACCACTCAGG
 1751 CATACTICC AGCATATATC TATCAGCTGG CAGGAGCTAC TTTGAAACCA
 1801 CCGGGCCCTA CTCTGACATT ATATACAACC CTTTCACAGA CAAAGGCACA
 1851 GGCAACATA TCTGGATAGA CTACCTCACA AARGAAGACA CCATTTTGT
 1901 GAAAAACAAA AGCAAATGCG AGATAATGGA CATGCCCTG TGGCGGCCT
 1951 GCACAGGATA CACAGAGTTT TGTGCAAAGT ATACAGGCGA CTCTGCCATT
 2001 ATCTACAATG CAAGAATACT CATAAGATGC CCATACACTG AGCCCATGTT
 2051 AATAGACCAC TCAGACCCAA ACAAAGCTT CGTTCCTAC TCATTTAACT
 2101 TTGGCAACGG AAAGATGCCG GGAGGAGCT CCAACGTGCC CATAAGAATG
 2151 AGAGCCAAGT GGTACGTGAA CATATTCCAC CAAAAAGAAG TATTAGAGAG
 2201 CATAGTACAG TCCGGACCGT TTGGGTACAA GGGCGACATA AGATCAGCTG
 2251 TACTAGCCAT GAAATACAGA TTTCACTGGA AGTGGGGCGG AAACCCATA
 2301 TCCAACACAG TCGTCAGGAA TCCCTGCTCC AACTCCAGCT CCTCCGCGGC
 2351 CCATAGAGGA CCTCGCAGCG TACAAGCGGT TGACCCGAAA TACAATACCC
 2401 CAGAGGTAC GTGGCACTCG TGGGACATTA GACGAGGACT CTTTGGCAAA
 2451 GCAGGTATTA AAAGAATGCA ACAGGAATCA GATGCTCTTT ACATTCCTCC
 2501 AGGACCAATC AAGAGACCTC GCAGGGACAC CAACGCCCAA GACCCAGAAG
 2551 AGCAAAACGA AAGCTCAGGT TTCAGAGTCC AGCAGGACT CCCGTGGGT
 2601 CACTCCAGCC AAGAGACGCA AAGCTCCCA GAAGAGACGG AGGCCAGGG
 2651 GTCGGTACAA GACCAACTAC TCCTCCAGCT CCGAGAGCAG CGAGTTCTCC
 2701 GACTCCAGCT CCAGCAACTC GCAACCCAAG TCCTCAAAGT CCAAGCAGGG
 2751 CACAGCCTAC ACCCCCTATT ATCTTCCCAA GCATAACAA AGCCTTTATG
 2801 TTTAGCCCC AGGGTCTTAA ACCCATAACG GGGTACAACG ACTGGCTAGA
 2851 AGAGTACACT CTTGCAAAAT TCTGGGACAG ACCCCCCAGA AAGTACACA
 2901 CAGACATACC CTTTACCCC TGGGCACCAA AACCCCAACA GCAAGTCAGG
 2951 GTGTCTTTA AACTCAACTT TCAATAAAAA TTCTAGGCCG TGGGAGTTT
 3001 ACTTGTGCGT GTCTGCTTCT TAAGGTCGCC AAGCACTCCG AGGCCAGCG
 3051 AGGAGTSCGA CCCCCCTCC GGTAGCAACG CCTTCGGAGC CGCGCGTAC
 3101 GCCTTCGGCT GCGCGCGGCA CCTCAGACCC CCCCTCCACC CGAAACGCTT
 3151 GCGCGTTTCG GACCTTCGGC GTCGGGGGGG TCGGGAGCTT TATTAACAG
 3201 ACTCCGAGTT GCCATTGGAC ACTGGAGCTG TGAATCAGTA ACGAAAGTGA
 3251 GTGGGGCCAG ACTTCGCCAT AGGGCCITTA TCTTCTCGCC ATTTGATAGT
 3301 GTCGGGGTC GCCGTAGGCT TCGGCCTCGT TTTTAGGCCT TCCGGACTAC
 3351 AAAAAATGGCG GTTTTAGTGA CGTCACGGC GCCATTTTAA GTAAGGCGGA
 3401 AGCAGCTCCA CTTTCTACA BAATGGCGGC GGAGCACTTC CGGCTTGCC
 3451 AAAATGGCGG GCAAGCTCTT CCGGTAAAG GGTGAGCAGC TACGTCACAA
 3501 GTCACCTGAC TGGGGAGGGG TCACAACCCG GAAGCCCTCC TCAGTCACGT
 3551 GGCTGTTTAC GTGGTTGCTA CGTCATCGGC GCCATCTTGT GTCGAAAAAT

```

3601 GCGCGACAAC TTCCGCTTTT TTAAAAAAG GCGCGAAAAA ACGGCGGGCG
3651 CGGCGCGCGC GCTGTGCGCG CGCGCCGGGG GGGCGCCAGC GCCCCCCCCC
3701 CCGCGCATGC GCGGGTCCCC CCCCCGCGG GGGGCTCCGC CCCCCGGCCC
3751 CCCCCCG

```

B

zpr9.6

zpr9.B1.6.seq (molécula subviral, tth25)
 293TT+9 (tth25smfr3) PCR
 Tr11.8 C/6
 En pCR2.1,ONESHOT Topo10F'

```

zpr9.B1.6.seq Longitud: 621
 1 CCGCAGCGAG AACGCCACGG AGGGAGATCC TCGCGTCCCG AGGGCGGGTG
51 CCGGAGGTGA GTTACACAC CGCAGTCAAG GGGCAATTTC GGCTCGGGAC
101 TGGCCGGGCC CCGGGCAAGG CTCTTAAAAA ATGCGTTTTC GCAGGGTTGC
151 CCAGAAAAGG AAAGTGCTTT TGCAAAGTGT GCCAGCTGCA AAGAAGGCTA
201 GCGGGCTTCT AGGTATGTGG CAGCCCCCA CGCACAATGT CCCGGGCATC
251 GAGAGAAACT GGTACGAGAG CTGTTTAGA TCCCACGCTG CTGTTTGTGG
301 CTGTGGCGAT TTGTGGCC ATCTTAATCA TCTGGCAACT ACTCTGGGTC
351 GTCCTCCGCG TCCTGGGCC CCAGGCGGAC CCCGCACGCC GCAAATAAGA
401 AACCTGCCAG CGTCCCGGC GCCCCAGGGC GAGCCCGGTG ACAGAGCGCC
451 ATGGCATGGG GCTTCTGGG CCGACGCCGC CGGTGGAGAC GATGGAGAGC
501 GCGGCGAGA CGGTGGAGAC CCCGAGGCC AACAGAGTC TACGTCTGTC
551 ACTTCCTGGG CATGGTCTAC ATCATAATAT AAGAACGTGC ACTTCCGAAT
601 GGCTGAGTTT TCCACGCCG T

```

Figura 13

A

ttrh215
 rheu.cd.215.seq
 Genomiphi tth4-cebador
 aso. 78°+ . taq polimerasa adicional (progr.tth4*78)
 vector:TA-cloning per2.1 invitrogen

rheu.cd.215rp.seq Longitud : 3758

```

1 AAAGTACGTC ACTAACCACG TGACTCCCAC AGGCCAACCA CAGTCTACGT
51 CGTGCATTTC CTGGGCATGG TCTACATCAT AATATAAGAA GCGCACCTC
101 CGAATGGCTG AGTTTCCAC GCCCGCCGC AGCGAGAACG CCACGGAGGG
151 AGATCCTCGC GTCCCGAGGG CGGGTGCCGG AGGTGAGTTT ACACACCGCA
201 GTC AAGGGGC AATTCGGGCT CGGGACTGGC CGGGCCCTGG GCAAGGCTCT
251 TAAAAAATGC GCTTTCGACG GGTGCGGAG AAAAGGAAAG TGCTTCTGCA
301 AACTCTGCGA GCTGCAAAGC AGGCTAGGCG GCTTCTAGGT ATGTGGCAGC
351 CCCCCGCGCA CAATGTCCC GGCATCGAGA GAAACTGGTA CGAGAGCTGC
401 TTCAGGTCTC ACGCTGCTGT TTGTGGCTGT GCGACTTTG TTGGCCATAT
451 TAATCATTTG GCAACTACTC TGGGTCGTCC TCCGCGTCTT GGGCCCCCAG
501 GCGGACCCCG CACGCCGCAA ATAAGAAACC TGCCAGCGCT CCCGGCGCCC
551 CAGGGCGAGC CCGGTGACAG AGCGCCATGG CGTGGGGTTT CTGGGGCCGA
601 CGCCCGCGGT GGAGACGGTG GAGAGCCCGG CGCAGACGGT GGAGACCCCG
651 GAGACGTAGG AGACACGCC CTGCTCGCCG CTTTCGAGCT CGTCGAAGAG
701 TAAGGAGACG CGGGGGGAGG TGGCGCAGAG GCTACAGAAA ATGGCGACGG
751 GGCAGACGCA GACGGACTCA CAGAAAAAAG ATAATTATTA AACAGTGGCA
801 ACCAAACTTT ATTAGACGCT GCTACATAAT AGGATGCCTA CCTCTCGTTT
851 TCTGTGGCGA AAATACAACC GCCCAGAACT ATGCCACTCA CTCAGACGAT
901 ATGATAAGCA AAGGACCGTA CGGGGGGGGG ATGACTRCCA CGAAATTCAC
951 TCTGAGAATA CTGTACGACG AGTTTACCAG GTTTATGAAC TTTTGGACTG
1001 TCAGTAACGA AGACCTAGAC CTGTGTAGAT ACGTGGGCTG CAAACTGATA
1051 TTTTTTAAAC ACCCCACGGT GGACTTTATG GTACAGATAA ACACTCAGCC
1101 TCCTTTCTTA GACACAAGCC TCACCGCGGC CAGCATACAC CCGGGCATCA
1151 TGATGCTCAG CAAGAGACGC ATATTAATAC CCTCTCTAAA GACCCGGCCG
1201 AGCAGAAAAC ACAGGGTGGT CGTCAGGGTG GGCGCCCAA GACTTTTTCA
1251 GGACAAGTGG TACCCCAAGT CAGACCTATG TGACACAGTT CTGCTTTCCA
1301 TATTTGCAAC CGCCCGGAC TTGCAATATC CGTTCGGCTC ACCACTAAT
1351 GACAACCCTT GCCTCAACTT CCAGATCCTG GGGCCCCAGT ACAAAAACA
1401 CCTTAGTATT AGTCCACTA TGGATGATAC TAACAACAG CACTATAACA
1451 GCAACTTATT TAATAAAACT GCACATACA ACACCTTTCA AACCRTAGCC
    
```

ES 2 657 705 T3

1501 CGGCTTAAAG AGACAGGACA AACTGCAAAC ATTAGTCCAA GTTGGAGTGA
 1551 AGTACAAAAC ACAAACACTAC TAGATCACAC AGGTGCTAAT GCAACTGCCA
 1601 GCAGAGACAC TTGGTACAAG GGAAACACAT ACAATGACTA CATACAACAG
 1651 TTAGCAGAGA AAACAGAGA AAGGTTTAAA AAAGCAACAA TGTCAGCACT
 1701 ACCAACTAC CCCACAATA TGTCCACAGA CTTATACGAA TACCACTCAG
 1751 GCATATACTC CAGCATATTT CTATCAGCTG GCAGGAGCTA CTTTGAACC
 1801 ACTGGGGCCT ACTCTGACAT TATATACAAC CCTTTGACAG ACAAAGGCAC
 1851 AGGCAACATA ATCTGGATAG ACTACCTTAC AAAAGACGAC ACAATCTTTG
 1901 TAAAAAACAA AAGCAAATGT GAGATAATGG ACATGCCCTT GTGGCGGGCC
 1951 GGCACAGGAT ACACAGAGTT TTGTGCAAAG TACACAGGAG ACTCTGCCAT
 2001 TATTTACAAT GCCAGAATAC TCATAAGATG CCCATACACT GAACCCATGC
 2051 TAATAGACCA CTCAGACCCA AACAAAGGCT TTGTACCCTA CTCATTTAAC
 2101 TTTGGCAACG GAAAGATGCC GGGAGGCAGC TCCAACGTGC CCATAAGANT
 2151 GAGAGCCAAG TGGTACGTAA ACATATCCA CCAAAAAGAA GTATTGGAGA
 2201 GCATAGTACA GTCGGACCG TTCGGGTACA GGGGCGACAT AAAATCAGCT
 2251 GTACTGTCCA TGAANTACAG ATTTCACTGG AAATGGGGCG GAAACCCAT
 2301 ATCCAACAG GTCGTCAGGA ATCCCTGCTC CAACTCCAGC ACCTCCGGCG
 2351 CCCATAGAGG ACCTCGCAGC GTACAAGCGG TTGACCCGAA ATACAATACC
 2401 CCAGAAGTCA CTTGGCACTC GTGGGACATC AGACGAGGAC TCTTTGGCAA
 2451 AGCAGGTATT AAAAGAATGC AACAGAATC AGATCCTCTT TACGTTCCTG
 2501 CAGGACCACT CAAGGGCCT CGCAGAGACA CCAACGCCA AGACCCGGAA
 2551 AAGCAAAACG AAAGCTCAGC TTTCGGAGTC CAGCAGCGAC TCCCGTGGGT
 2601 CCCTCCAGC CAAGAGACGC AAAGCTCCA AGAAGAGACG CAGGCGCAGG
 2651 GGTGCTACA AGACCAACTA CTCCTCCAGC TCCGAGAGCA GCGAGTACTC
 2701 CGACTCCAGC TCCAACAAC CGCACCCCAA GTCCCTCAAAG TTCAAGCAGG
 2751 ACACAGCCTA CACCCCTAT TATCCTCCA AGCATAAACA AAGCCTATAT
 2801 GTTTGAACCC CAGGGTCCCT AACCCATACA GGGGTACAAC GATTGGCTAG
 2851 AGGAGTACAC TAGTTGCAAG TTCCGGGACA GACCCCGAG AATGCTACAC
 2901 ACAGACTTAC CCTTTTACCC CTGGGCACCA AAACCCCAAG ACCAAGTCAG
 2951 GGTAACTTT ARACTCAACT TTCATAAAA ATTCTAGGCC GTGGGACTTT
 3001 CACTTGTCCG TGCTGCTTC TTAAGGTCGC CAAGCACTCC GAGCGTCAGC
 3051 GAGGAGTCCG ACCCCCCCC TCGGTAGCAA CGCCTTCGGA GCCGCGCGCT
 3101 ACGCCTTCGG CTGCGCGCG CACCTCAGAC CCCCCCTCA CCCGAAACGC
 3151 TTGCGCGTTT CGGACCTTCG GCGTCGGGGG GGTGCGGAGC TTTATTAAC
 3201 AGACTCCGAG TTGCCATGG AACTGGAGC TGTGAATCAG TAACGAAAGT
 3251 GAGTGGGGCC AGACTTCGCC ATAGGGCCTT TATCTTCTCG CCATTGGATA
 3301 GTGTCGGGG TTGCCGTAGG CTTCCGGCTC GTTTTAGGC CTTCCGGACT
 3351 ACAAATATGG CGGATTTGT GACGTCACGG CCGCCATTTT AAGTAAGGGC
 3401 GAAGCAGCTC CACCCCTCA CATAATGGG GCGGAGCACT CCCGGCTTGC

ES 2 657 705 T3

```
3451 CCAAAATGGC GGGCAAGCTC TTCCGGGTCA AAGGTTGGCA GCTACGTCAC
3501 AAGTCACCTG ACTGGGGAGG AGTTACATCC CGGAAGTTCT CCTCGGTCCAC
3551 GTGACTGTAC ACGTACTGTC TACGTCATTG ACGCCATCTT GTGTCACAAA
3601 ATGGCGGTGC ACTTCCGCTT TTTTAAAAAA AGGCGCGAAA AAACGGCGGC
3651 GCGCGCGCGC GCGCTGCGCG CGCGCGCCGG GGGGGCGCCA GCGCCCCCC
3701 CCCCAGCAT GCACGGGTCC CCCCCCAC GGGGGCTCC GCCCCCGGC
3751 CCCCCC
```

B

zpr12.24

zpr12.24.seq : (molécula subvívica , ttrh215)

Longitud: 642

```
1 CAGCGAGAAC GCCACGGAGG GAGATCCTCG CGTCCCAGG GCGGGTGCCG
51 GAGGTGAGTT TACACACCGC AGTCAAGGGG CAATTCGGGC TCGGGACTGG
101 CCGGGCCCCG GGCAAGGCTC TTAAAAAATG CGCTTTCGCA GGGTTGCTGA
151 GAAAAGGAAA GTGCTTCTGC AACTGTGCG AGCTACACAG AAGACTAGGC
201 GGCTTCTAAG CCGCCCACAG GGGCATGTCT ACATGCTTCC GCAGCGAGAA
251 CGCCACGGAG GGAGATCCTC GCGTCCCAGG GCGGGGTGCC GGAGGTGAGT
301 TTACACACCG CAGTCAAAGG GCAATTCGGG CTCGGGACTG GCCGGGCCCC
351 GGGCAAGGCT CTTAAAAAAT GCGCTTTCGC GGGTTGCTG AGAAAAGGAA
401 AGTGCTTCTG CAAACTGTGC GAGCTACACA GAAGACTAGG CGGCTTCTAG
451 GTATGTGGCA GCCCCCGTG CACAATGTCC CCGGCATCTT ATTAGTACTC
501 TGGCGTTGTA GATAATGGCA GAGTCTCCAG TGTACTTTGC ACAGAACTCT
551 GTGTATCCTG TGCAGGCCGC CCACAGGGGC ATGCTACAT CATAATATAA
601 TAAGGCGCAC TTCCGAATGG CTGAGTTTTC CAGCCCCGTC CG
```

Figura 14

Marcos de lectura abierta de 71 nt (HCR) :

```

zyb2.1.pep  RVPKVS LHTA VKGQFGLGTG RAM
zyb9.1.pep  RVPKVS LHTA VKGQFGLGTG RAM
zkb69.1.pep RVPEVSLHTA VKGQFGLGTG RAM

zyb2.3.pep  GAEGEFTHRS QGAI RARDWP GHG
zyb9.3.pep  GAEGEFTHRS QGAI RARDWP GYG
zkb5.3.pep  GAVGEFTHRS QGAI RARDWP GYG
zkb69.3.pep GAGGEFTHRS QGAI RARDWP GYG
    
```

zyb2.1.pep
 nucléotidos 1-71
 Longitud: 23aa

RVPKVS LHTA VKGQFGLGTG RAM

```

BlastP2 de: zyb2.1.pep
comparado con la base de datos: uniprot ..
>>>sptrembl:Q9WSW0_9VIRU Q9sw0 SubName: Full=ORF2; Flags: Fragm... 49 1e-04
>>>sptrembl:Q9WB09_9VIRU Q9wb09 SubName: Full=ORF2; Flags: Fragm... 48 4e-04
>>>sptrembl:Q9WSW2_9VIRU Q9sw2 SubName: Full=ORF2; Flags: Fragm... 48 4e-04
>>>sptrembl:Q9WSX0_9VIRU Q9wsx0 SubName: Full=ORF2; Flags: Fragm... 47 4e-04
>>>sptrembl:Q9WB10_9VIRU Q9wb10 SubName: Full=ORF2; Flags: Fragm... 47 5e-04
>>>sptrembl:Q9WAZ2_9VIRU Q9waz2 SubName: Full=ORF2; Flags: Fragm... 47 5e-04
>>>sptrembl:O70807_9VIRU O70807 SubName: Full=Putative uncharact... 47 6e-04
>>>sptrembl:Q9WAY4_9VIRU Q9way4 SubName: Full=ORF2; Flags: Fragm... 47 7e-04
>>>sptrembl:Q9WB02_9VIRU Q9wb02 SubName: Full=ORF2; Flags: Fragm... 47 8e-04
>>>sptrembl:Q9WSW4_9VIRU Q9sw4 SubName: Full=ORF2; Flags: Fragm... 46 0.001
>>>sptrembl:Q9WB12_9VIRU Q9wb12 SubName: Full=ORF2; Flags: Fragm... 46 0.001
>>>sptrembl:Q9WSW6_9VIRU Q9sw6 SubName: Full=ORF2; Flags: Fragm... 45 0.002
>>>sptrembl:B3FWR6_9VIRU B3fwr6 SubName: Full=ORF2; Flags: Fragm... 36 0.94

>>>>sptrembl:Q9WSW0_9VIRU Q9sw0 SubName: Full=ORF2; Flags:
Fragmento;. 2/2010
Longitud = 204
Puntuación = 49.3 bits (116), Esperada = 1e-04, Método: ajuste de matriz composicional
Identities = 23/23 (100%), Positivos = 23/23 (100%)
Query: 1 RVPKVS LHTAVKQFGLGTGRAM 23
RVPKVS LHTAVKQFGLGTGRAM
Sbjct: 27 RVPKVS LHTAVKQFGLGTGRAM 49

>>>>sptrembl:Q9WB09_9VIRU Q9wb09 SubName: Full=ORF2; Flags:
Fragmento;. 2/2010
Longitud = 138
Puntuación = 47.8 bits (112), Esperada = 4e-04, Método: ajuste de matriz composicional
Identities = 22/23 (95%), Positivos = 22/23 (95%)
Query: 1 RVPKVS LHTAVKQFGLGTGRAM 23
RVPKVS LHT VKGQFGLGTGRAM
Sbjct: 14 RVPKVS LHTEVKQFGLGTGRAM 36

ID Q9WSW0_9VIRU Sin revisar; 204 AA.
    
```

```

AC Q9WSW0;
DT 01-NOV-1999, integrado en UniProtKB/TrEMBL.
DT 01-NOV-1999, versión de secuencia 1.
DT 09-FEB-2010, versión de entrada 23.
DE SubName: Full=ORF2;
DE Flags: Fragmento;
OS Torque teno virus.
OC Virus; virus de ADNss; Anelloviridae; Anelloviridae sin clasificar.
OX NCBI_TaxID=68887;
RN [1]
RP SECUENCIA DE NUCLEÓTIDOS
RC CEPA =KC205/1-12G; TEJIDO=Suerdo;
RX MEDLINE=99335592; PubMed=10405352; DOI=10.1006/viro.1999.9797;
RA Hijikata M., Takahashi K., Mishiro S.;
RT "Complete circular DNA genome of a TT virus variant (isolate name
RT SANBAN) and 44 partial ORF2 sequences implicating a great degree of
RT diversity beyond genotypes.";
RL Virology 260:17-22(1999).
CC -----
CC Derechos de autor de UniProt Consortium, véase http://www.uniprot.org/terms
CC Distribuido bajo la licencia de Creative Commons Attribution-NoDerivs
CC -----
DR EMBL; AB024383; BAA77450.2; -; Genomic DNA.
DR InterPro; IPR004118; Gyrovir_VP2/TT_ORF2.
DR Pfam; PF02957; TT_ORF2; 1.
PE 4: Previsto;
FT NON TER 1 1
SQ SECUENCIA 204 AA; 21953 MW; 6352C96D2AC0DF21 CRC64;
CTSEWLSFPR PSAAAXPRRV IPASRWVRVPK VSLHTAVKGQ FGLGTGRAMG KALKVFILKM
HFSRISRSKR KVLLPALPAP PPPRQLLMWQ PPIQNGTQLD RHWFESVWRS HAAYCGCGDG
VGHLLQHLAAN LGRPPHPQPP REQHPPQIRG LPALPAPPSN RNSWPGTGGD AAGEQAGGSR
GAGDGGDGEL ADDDLXDAAA LVEE

ID Q9WB09_9VIRU Sin revisar ; 138 AA.
AC Q9WB09;
DT 01-NOV-1999, integrado en UniProtKB/TrEMBL.
DT 01-NOV-1999, versión de secuencia 1.
DT 09-FEB-2010, versión de entrada 22.
DE SubName: Full=ORF2;
DE Flags: Fragmento;
OS Torque teno virus.
OC Virus; virus de ADNss; Anelloviridae; Anelloviridae sin clasificar.
OX NCBI_TaxID=68887;
RN [1]
RP SECUENCIA DE NUCLEÓTIDOS
RC TEJIDO=Suerdo;
RX MEDLINE=99335592; PubMed=10405352; DOI=10.1006/viro.1999.9797;
RA Hijikata M., Takahashi K., Mishiro S.;
RT "Complete circular DNA genome of a TT virus variant (isolate name
RT SANBAN) and 44 partial ORF2 sequences implicating a great degree of
RT diversity beyond genotypes.";
RL Virology 260:17-22(1999).
CC -----
CC Derechos de autor de UniProt Consortium, véase http://www.uniprot.org/terms
CC Distribuido bajo la licencia de Creative Commons Attribution-NoDerivs
CC -----
DR EMBL; AB024376; BAA77443.1; -; Genomic DNA.
DR InterPro; IPR004118; Gyrovir_VP2/TT_ORF2.
DR Pfam; PF02957; TT_ORF2; 1.
PE 4: Previsto;
FT NON TER 1 1
FT NON TER 138 138
SQ SECUENCIA 138 AA; 15494 MW; 2DF27B3A4F0CA641 CRC64;

```

ES 2 657 705 T3

AVKPRREISA SRGRVPKVS L HTEVKGQFGL GTGRAMGKAL KKS MFIGRH Y RKKRALSLCA
 VRTTKKACKL LIVMWT PPRN DQYLNWQWY SSVLSSHAAM CGCPDAIAHL SHLAFVFRAP
 QNPPPPGQQR NLPLRRLP

zyb2.3.pep

nucleótidos 3-71

Longitud: 23aa

GAEGEFTHRS QGAIRARDWP GHG

BlastP2 de: zyb2.3.pep desde: 1 a: 23

comparado con la base de datos: **uniprot** ..

>>>sptrembl:Q98Y39 9VIRU Q98y39	SubName: Full=ORF2;. 2/2010	50	7e-05
>>>sptrembl:Q9WAY7 9VIRU Q9way7	SubName: Full=ORF2; Flags: Fragm...	50	7e-05
>>>sptrembl:Q786D4 9VIRU Q786d4	SubName: Full=ORF2;. 2/2010	50	9e-05
>>>sptrembl:Q70738 9VIRU Q70738	SubName: Full=ORF2, ORF1 genes;...	50	9e-05
>>>sptrembl:Q9JG33 9VIRU Q9jg33	SubName: Full=Putative uncharact...	50	9e-05
>>>sptrembl:Q90363 9VIRU Q90363	SubName: Full=ORF2 protein;. 2/2010	50	9e-05
>>>sptrembl:Q9WFY6 9VIRU Q9wfy6	SubName: Full=ORF2;. 2/2010	50	1e-04
>>>sptrembl:Q9JGT0 9VIRU Q9jgt0	SubName: Full=PORF2a;. 12/2009	50	1e-04
>>>sptrembl:Q9YKL2 9VIRU Q9ykl2	SubName: Full=ORF2 protein;. 2/2010	50	1e-04
>>>sptrembl:Q9JGS7 9VIRU Q9jgs7	SubName: Full=PORF2a;. 12/2009	49	1e-04
>>>sptrembl:Q9DYC0 9VIRU Q9dyc0	SubName: Full=Putative uncharact...	49	1e-04
>>>sptrembl:Q9W7S4 9VIRU Q9w7s4	SubName: Full=Putative uncharact...	49	2e-04
>>>sptrembl:Q77S01 9VIRU Q77s01	SubName: Full=Putative uncharact...	49	2e-04
>>>sptrembl:Q9JGT3 9VIRU Q9jgt3	SubName: Full=PORF2a;. 12/2009	46	0.001
>>>sptrembl:Q9JGS4 9VIRU Q9jgs4	SubName: Full=PORF2a;. 12/2009	46	0.001
>>>sptrembl:Q9YR02 9VIRU Q9yr02	SubName: Full=Putative uncharact...	45	0.003
>>>sptrembl:B2YFW4 9VIRU B2yfw4	SubName: Full=ORF2; Flags: Fragm...	34	3.7

>>>sptrembl:Q98Y39 9VIRU Q98y39 SubName: Full=ORF2;. 2/2010

Longitud = 202

Puntuación = 50.1 bits (118), Esperado = 7e-05, Método: ajuste de matriz composicional

Identidades: = 22/23 (95%), Positivos = 23/23 (100%)

Query: 1 GAEGEFTHRSQGAIRARDWPGHG 23

GAEGEFTHRSQGAIRARDWPG+G

Sbjct: 24 GAEGEFTHRSQGAIRARDWPGYG 46

Q98Y39_9VIRU Sin revisar ; 202 AA.

AC Q98Y39;

DT 01-JUN-2001, : integrado en UniProtKB/TrEMBL.

DT 01-JUN-2001, versión de secuencia 1.

DT 09-FEB-2010, versión de entrada 17.

DE SubName: Full=ORF2;

OS Torque teno virus.

OC Virus; virus de ADNss; Anelloviridae; Anelloviridae sin clasificar.

OX NCBI_TaxID=68887;

RN [1]

RP SECUENCIA DE NUCLEÓTIDOS.

RC CEPA =TWH;

RA He H.-T., Luo K.-X., Xiao H., Liu D.-X.;

RT "Complete circular genome of TT virus isolated from feces of a

RT hepatitis patient.;"

RL Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.

CC

CC Derechos de autor de UniProt Consortium, véase <http://www.uniprot.org/terms>

CC Distribuido bajo la licencia de Creative Commons Attribution-NoDerivs

CC

DR EMBL; AF351132; AAK29446.1; -; Genomic DNA.

DR InterPro; IPR004118; Gyrovir_VP2/TT_ORF2.

DR InterPro; IPR013267; TTV_ORF2a.

ES 2 657 705 T3

```

DR Pfam; PF02957; TT_ORF2; 1.
DR Pfam; PF08197; TT_ORF2a; 1.
PE 4; Previsto;
SQ SECUENCIA 202 AA; 21437 MW; 105B9ED104956EDE CRC64;
MAEFSTPVRS GEATEGDHRV PRAGAEGEFT HRSQGAIRAR DWPGYGQSE KSMFIGRHYR
KKRALSLCAV RTTKKACKLL IVMWTPPRND QQYLNWQWYS SVLSSHASMC GCPDAVAHLI
NLASVLRAPQ NPPPPGPQRN LPLRRLPALP AAPEAPGDRA PWPAGGAEG ENGGAGGDAD
HGGAAGGPED ANLLDAVAAA ET
//

```

zyb9.1.pep

nucleótidos 1-71
Longitud: 23aa

RVPKVS~~L~~HTA VKGQFGLGTG RAM

BlastP2 de: zyb9.1.pep desde: 1 a: 23
comparado con la base de datos: **uniprot** ..

>>>sptrembl:Q9WSW0_9VIRU	Q9wsw0	SubName: Full=ORF2; Flags: Fragm...	49	1e-04
>>>sptrembl:Q9WB09_9VIRU	Q9wb09	SubName: Full=ORF2; Flags: Fragm...	48	4e-04
>>>sptrembl:Q9WSW2_9VIRU	Q9wsw2	SubName: Full=ORF2; Flags: Fragm...	48	4e-04
>>>sptrembl:Q9WSX0_9VIRU	Q9wsx0	SubName: Full=ORF2; Flags: Fragm...	47	4e-04
>>>sptrembl:Q9WB10_9VIRU	Q9wb10	SubName: Full=ORF2; Flags: Fragm...	47	5e-04
>>>sptrembl:Q9WAZ2_9VIRU	Q9waz2	SubName: Full=ORF2; Flags: Fragm...	47	5e-04
>>>sptrembl:O70807_9VIRU	O70807	SubName: Full=Putative uncharact...	47	6e-04
>>>sptrembl:Q9WAY4_9VIRU	Q9way4	SubName: Full=ORF2; Flags: Fragm...	47	7e-04
>>>sptrembl:Q9WB02_9VIRU	Q9wb02	SubName: Full=ORF2; Flags: Fragm...	47	8e-04
>>>sptrembl:Q9WSW4_9VIRU	Q9wsw4	SubName: Full=ORF2; Flags: Fragm...	46	0.001
>>>sptrembl:Q9WB12_9VIRU	Q9wb12	SubName: Full=ORF2; Flags: Fragm...	46	0.001
>>>sptrembl:Q9WSW6_9VIRU	Q9wsw6	SubName: Full=ORF2; Flags: Fragm...	45	0.002
>>>sptrembl:B3FWR6_9VIRU	B3fwr6	SubName: Full=ORF2; Flags: Fragm...	36	0.94

>>>>sptrembl:Q9WSW0_9VIRU Q9wsw0 SubName: Full=ORF2; Flags:
Fragmento ;. 2/2010
Longitud = 204
Puntuación= 49.3 bits (116), Esperado = 1e-04, Método: ajuste de matriz composicional.
Identidades = 23/23 (100%), Positivos = 23/23 (100%)
Query: 1 RVPKVS~~L~~HTAVKQFGLGTGRAM 23
RVPKVS~~L~~HTAVKQFGLGTGRAM
Sbjct: 27 RVPKVS~~L~~HTAVKQFGLGTGRAM 49

>>>>sptrembl:Q9WB09_9VIRU Q9wb09 SubName: Full=ORF2; Flags:
Fragmento ;. 2/2010
Longitud = 138
Puntuación= 47.8 bits (112), Esperado = 4e-04, Método: ajuste de matriz composicional.
Identidades = 22/23 (95%), Positivos = 22/23 (95%)
Query: 1 RVPKVS~~L~~HTAVKQFGLGTGRAM 23
RVPKVS~~L~~HTA VKGQFGLGTGRAM
Sbjct: 14 RVPKVS~~L~~HTAVKQFGLGTGRAM 36

Q9WSW0_9VIRU Sin revisar ; 204 AA.
AC Q9WSW0;
DT 01-NOV-1999, integrado en UniProtKB/TrEMBL.
DT 01-NOV-1999, versión de secuencia 1.
DT 09-FEB-2010, versión de entrada 23.
DE SubName: Full=ORF2;
DE Flags: Fragmento ;

OS Torque teno virus.
 OC Virus; virus de ADNss; Anelloviridae; Anelloviridae sin clasificar.
 OX NCBI_TaxID=68887;
 RN [1]
 RP SECUENCIA DE NUCLEÓTIDOS.
 RC CEPA =KC205/1-12G; TEJIDO=Sueros;
 RX MEDLINE=99335592; PubMed=10405352; DOI=10.1006/viro.1999.9797;
 RA Hijikata M., Takahashi K., Mishiro S.;
 RT "Complete circular DNA genome of a TT virus variant (isolate name
 RT SANBAN) and 44 partial ORF2 sequences implicating a great degree of
 RT diversity beyond genotypes.";
 RL Virology 260:17-22(1999).
 CC -----
 CC Derechos de autor de UniProt Consortium, véase <http://www.uniprot.org/terms>
 CC Distribuido bajo la licencia de Creative Commons Attribution-NoDerivs
 CC -----
 DR EMBL; AB024383; BAA77450.2; -; Genomic DNA.
 DR InterPro; IPR004118; Gyrovir_VP2/TT_ORF2.
 DR Pfam; PF02957; TT_ORF2; 1.
 PE 4: Previsto;
 FT NON TER 1 1
 SQ SECUENCIA 204 AA; 21953 MW; 6352C96D2AC0DF21 CRC64;
 CTSEWLSFPR PSAAAXPRRV IPASRWRVVK VSLHTAVKGG FGLGTGRAMG KALKVFILKM
 HFSRISRSKR KVLLPALPAP PPRQLLMWQ PPIQNGTQLD RHWFEVWRS HAAYCGCGDC
 VGHLLQHLAAN LGRPPHPQPP REQHPPIRG LPALPAPPNS RNSWPGTGGD AAGEQAGGSR
 GAGDGGDDEL ADDDLXDAAA LVEE

Q9WB09 GVIRU

Sin revisar;

138 AA.

AC Q9WB09;
 DT 01-NOV-1999, integrado en UniProtKB/TrEMBL.
 DT 01-NOV-1999, versión de secuencia 1.
 DT 09-FEB-2010, versión de entrada 22.
 DE SubName: Full=ORF2;
 DE Flags: Fragmento;
 OS Torque teno virus.
 OC Virus; virus de ADNss; Anelloviridae; Anelloviridae sin clasificar.
 OX NCBI_TaxID=68887;
 RN [1]
 RP SECUENCIA DE NUCLEÓTIDOS.
 RC TEJIDO=Sueros;
 RX MEDLINE=99335592; PubMed=10405352; DOI=10.1006/viro.1999.9797;
 RA Hijikata M., Takahashi K., Mishiro S.;
 RT "Complete circular DNA genome of a TT virus variant (isolate name
 RT SANBAN) and 44 partial ORF2 sequences implicating a great degree of
 RT diversity beyond genotypes.";
 RL Virology 260:17-22(1999).
 CC -----
 CC Derechos de autor de UniProt Consortium, véase <http://www.uniprot.org/terms>
 CC Distribuido bajo la licencia de Creative Commons Attribution-NoDerivs
 CC -----
 DR EMBL; AB024376; BAA77443.1; -; Genomic DNA.
 DR InterPro; IPR004118; Gyrovir_VP2/TT_ORF2.
 DR Pfam; PF02957; TT_ORF2; 1.
 PE 4: Previsto;
 FT NON TER 1 1
 FT NON TER 138 138
 SQ SECUENCIA 138 AA; 15494 MW; 2DF27B3A4F0CA641 CRC64;
 AVKPRREISA SRGRVPKVS LHTAVKGG FGLGTGRAMG KALKVFILKM RKKRALS LCA
 VRTTKKACKL LIVMWT PPRN DQYLNWQWY SSVLSSHAAM CGCPDAIAHL SHLAFVFRAP
 QNPPPPGPQR NLPLRRLP

zyb9.3.pep

nucleótidos 3-71
Longitud: 23aa

GAEGEFTHRS QGAIRARDWP GYG

BlastP2 de: zyb9.3.pep desde: 1 a: 23
comparado con la base de datos: **uniprot** ..

>>>sptrembl:Q9WAY7_9VIRU	Q9way7	SubName: Full=ORF2; Flags: Fragm...	52	2e-05
>>>sptrembl:Q98Y39_9VIRU	Q98y39	SubName: Full=ORF2;. 2/2010	52	2e-05
>>>sptrembl:Q9WFY6_9VIRU	Q9wfy6	SubName: Full=ORF2;. 2/2010	52	3e-05
>>>sptrembl:Q9JGT0_9VIRU	Q9jgt0	SubName: Full=PORF2a;. 12/2009	52	3e-05
>>>sptrembl:Q786D4_9VIRU	Q786d4	SubName: Full=ORF2;. 2/2010	52	3e-05
>>>sptrembl:Q70738_9VIRU	Q70738	SubName: Full=ORF2, ORF1 genes;...	52	3e-05
>>>sptrembl:Q9JG33_9VIRU	Q9jg33	SubName: Full=Putative uncharact...	52	3e-05
>>>sptrembl:Q90363_9VIRU	Q90363	SubName: Full=ORF2 protein;. 2/2010	52	3e-05
>>>sptrembl:Q9YKL2_9VIRU	Q9ykl2	SubName: Full=ORF2 protein;. 2/2010	51	3e-05
>>>sptrembl:Q9JGS7_9VIRU	Q9jgs7	SubName: Full=PORF2a;. 12/2009	51	3e-05
>>>sptrembl:Q9DYC0_9VIRU	Q9dyc0	SubName: Full=Putative uncharact...	51	4e-05
>>>sptrembl:Q9W7S4_9VIRU	Q9w7s4	SubName: Full=Putative uncharact...	50	5e-05
>>>sptrembl:Q77S01_9VIRU	Q77s01	SubName: Full=Putative uncharact...	50	5e-05
>>>sptrembl:Q9JGS4_9VIRU	Q9jgs4	SubName: Full=PORF2a;. 12/2009	48	3e-04
>>>sptrembl:Q9JGT3_9VIRU	Q9jgt3	SubName: Full=PORF2a;. 12/2009	48	3e-04
>>>sptrembl:Q9YR02_9VIRU	Q9yr02	SubName: Full=Putative uncharact...	47	7e-04
>>>sptrembl:B2YFW4_9VIRU	B2yfw4	SubName: Full=ORF2; Flags: Fragm...	35	3.2

>>>>sptrembl:Q9WAY7_9VIRU Q9way7 SubName: Full=ORF2; Flags:
Fragmento 2010
Longitud = 138
Puntuación= 52.0 bits (123), Esperado = 2e-05, Método: ajuste de matriz composicional.
Identidades = 23/23 (100%), Positivos = 23/23 (100%)
Query: 1 GAEGEFTHRSQGAIRARDWPGYG 23
GAEGEFTHRSQGAIRARDWPGYG
Sbjct: 15 GAEGEFTHRSQGAIRARDWPGYG 37

Q9WAY7_9VIRU Sin revisar.; 138 AA.
AC Q9WAY7;
DT 01-NOV-1999, integrado en UniProtKB/TrEMBL.
DT 01-NOV-1999, versión de secuencia 1.
DT 09-FEB-2010, versión de entrada 23.
DE SubName: Full=ORF2;
DE Flags: Fragmento ;
OS Torque teno virus.
OC Virus; virus de ADNss ; Anelloviridae; Anelloviridae sin clasificar.
OX NCBI_TaxID=68887;
RN [1]
RP SECUENCIA DE NUCLEÓTIDOS.
RC TEJIDO=Sueros;
RX MEDLINE=99335592; PubMed=10405352; DOI=10.1006/viro.1999.9797;
RA Hijikata M., Takahashi K., Mishiro S.;
RT "Complete circular DNA genome of a TT virus variant (isolate name
RT SANBAN) and 44 partial ORF2 sequences implicating a great degree of
RT diversity beyond genotypes.";
RL Virology 260:17-22 (1999).
CC -----
CC Derechos de autor de UniProt Consortium, véase <http://www.uniprot.org/terms>
CC Distribuido bajo la licencia de Creative Commons Attribution-NoDerivs
CC -----

DR EMBL; AB024348; BAA77415.1; -; Genomic DNA.
 DR InterPro; IPR004118; Gyrovir VP2/TT_ORF2.
 DR InterPro; IPR013267; TTV_ORF2a.
 DR Pfam; PF02957; TT_ORF2; 1.
 DR Pfam; PF08197; TT_ORF2a; 1.
 PE 4: Previsto;
 FT NON TER 1 1
 FT NON TER 138 138
 SQ SECUENCIA 138 AA; 15416 MW; 93F7D9685085141D CRC64;
 SGEATEGDLR VPRAGAEGEF THRSQGAIRA RDWPGYGQGS EKSMFIGRHY RKKRALSLCA
 VRTTKACKL LIVMWTPPRN DQYLNWQWY SSVLSSHAAM CGCPDAVAHF NHLAAVLRAP
 QNPPPPGPQR NLPLRRLP

zkb5.3.pep

nucleótidos 3-71
 Longitud: 23aa

GAVGEFTHRS QGAIRARDWP GYG

BlastP2 de: zkb5.3.pep
 comparado con la base de datos: **uniprot** ..

>>>sptrembl:Q98Y39_9VIRU	Q98y39	SubName: Full=ORF2;. 2/2010	49	1e-04
>>>sptrembl:Q9WAY7_9VIRU	Q9way7	SubName: Full=ORF2; Flags: Fragm...	49	1e-04
>>>sptrembl:Q9JGT0_9VIRU	Q9jgt0	SubName: Full=PORF2a;. 12/2009	49	1e-04
>>>sptrembl:Q90363_9VIRU	Q90363	SubName: Full=ORF2 protein;. 2/2010	49	1e-04
>>>sptrembl:Q9JG33_9VIRU	Q9jg33	SubName: Full=Putative uncharact...	49	1e-04
>>>sptrembl:Q786D4_9VIRU	Q786d4	SubName: Full=ORF2;. 2/2010	49	1e-04
>>>sptrembl:Q70738_9VIRU	Q70738	SubName: Full=ORF2, ORF1 genes;...	49	1e-04
>>>sptrembl:Q9YKL2_9VIRU	Q9ykl2	SubName: Full=ORF2 protein;. 2/2010	49	1e-04
>>>sptrembl:Q9Wfy6_9VIRU	Q9wfy6	SubName: Full=ORF2;. 2/2010	49	1e-04
>>>sptrembl:Q9JGS7_9VIRU	Q9jgs7	SubName: Full=PORF2a;. 12/2009	49	2e-04
>>>sptrembl:Q9DYC0_9VIRU	Q9dyc0	SubName: Full=Putative uncharact...	49	2e-04
>>>sptrembl:Q9W7S4_9VIRU	Q9w7s4	SubName: Full=Putative uncharact...	49	2e-04
>>>sptrembl:Q77S01_9VIRU	Q77s01	SubName: Full=Putative uncharact...	49	2e-04
>>>sptrembl:Q9JGS4_9VIRU	Q9jgs4	SubName: Full=PORF2a;. 12/2009	48	3e-04
>>>sptrembl:Q9JGT3_9VIRU	Q9jgt3	SubName: Full=PORF2a;. 12/2009	48	3e-04
>>>sptrembl:Q9YR02_9VIRU	Q9yr02	SubName: Full=Putative uncharact...	47	7e-04

>>>>sptrembl:Q98Y39_9VIRU Q98y39 SubName: Full=ORF2;. 2/2010
 Longitud = 202
 Puntuación = 49.3 bits (116), Esperado = 1e-04, Método: ajuste de matriz composicional.
 Identities = 22/23 (95%), Positivos = 22/23 (95%)
 Query: 1 GAVGEFTHRSQGAIRARDWPGYG 23
 GA GEFTHRSQGAIRARDWPGYG
 Sbjct: 24 GAEGEFTHRSQGAIRARDWPGYG 46

Q98Y39_9VIRU Sin revisar ; 202 AA.
 AC Q98Y39;
 DT 01-JUN-2001, integrado en UniProtKB/TrEMBL.
 DT 01-JUN-2001, versión de secuencia 1.
 DT 09-FEB-2010, versión de entrada 17.
 DE SubName: Full=ORF2;
 OS Torque teno virus.
 OC Virus; virus de ADNss ; Anelloviridae; Anelloviridae sin clasificar.
 OX NCBI_TaxID=68887;
 RN [1]
 RP SECUENCIA DE NUCLEÓTIDOS.
 RC CEPA =TWH;

RA He H.-T., Luo K.-X., Xiao H., Liu D.-X.;
 RT "Complete circular genome of TT virus isolated from feces of a
 RT hepatitis patient."
 RL Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
 CC -----
 CC Derechos de autor de UniProt Consortium, véase <http://www.uniprot.org/terms>
 CC Distribuido bajo la licencia de Creative Commons Attribution-NoDerivs
 CC -----
 DR EMBL; [AF351132](#); [AAK29446.1](#); -; Genomic_DNA.
 DR InterPro; [IPR004118](#); Gyrovir_VP2/TT_ORF2.
 DR InterPro; [IPR013267](#); TTV_ORF2a.
 DR Pfam; [PF02957](#); TT_ORF2; 1.
 DR Pfam; [PF08197](#); TT_ORF2a; 1.
 PE 4: Previsto ;
 SQ SECUENCIA 202 AA; 21437 MW; 105B9ED104956EDE CRC64;
 MAEFSTPVRS GEATEGDHRV PRAGAEGEFT HRSQGAIARR DWPYGGQSE KSMFIGRHYR
 KKRALSLCAV RTTKKACKLL IVMWTPPRND QQYLNWQWYS SVLSSHASMC GCPDAVAHLI
 NIASVLRAPQ NPPPPGPQRN LPLRRLPALP AAPEAPGDRA PWPAGGAEG ENGGAGGDAD
 HGGAGGPEP ANLLDAVAAA ET
 //

zkb69.1.pep

nucleótidos 1-71
 Longitud: 23aa

RVPEVSLHTA VKGQFGLGTG RAM

BlastP2 de: zkb69.1.pep desde: 1 a: 23
 comparado con la base de datos: uniprot

>>>sptrembl:Q9WAZ2_9VIRU	Q9waz2	SubName: Full=ORF2; Flags: Fragm...	<u>49</u>	1e-04
>>>sptrembl:Q9WB10_9VIRU	Q9wb10	SubName: Full=ORF2; Flags: Fragm...	<u>49</u>	1e-04
>>>sptrembl:Q70807_9VIRU	O70807	SubName: Full=Putative uncharact...	<u>49</u>	2e-04
>>>sptrembl:Q9WAY4_9VIRU	Q9way4	SubName: Full=ORF2; Flags: Fragm...	<u>49</u>	2e-04
>>>sptrembl:Q9WSW0_9VIRU	Q9wswo	SubName: Full=ORF2; Flags: Fragm...	<u>48</u>	3e-04
>>>sptrembl:Q9WB12_9VIRU	Q9wb12	SubName: Full=ORF2; Flags: Fragm...	<u>48</u>	4e-04
>>>sptrembl:Q9WSW2_9VIRU	Q9wswo	SubName: Full=ORF2; Flags: Fragm...	<u>48</u>	4e-04
>>>sptrembl:Q9WSW4_9VIRU	Q9wswo	SubName: Full=ORF2; Flags: Fragm...	<u>47</u>	7e-04
>>>sptrembl:Q9WB09_9VIRU	Q9wb09	SubName: Full=ORF2; Flags: Fragm...	<u>47</u>	9e-04
>>>sptrembl:Q9WSX0_9VIRU	Q9wsx0	SubName: Full=ORF2; Flags: Fragm...	<u>46</u>	0.001
>>>sptrembl:Q9WB02_9VIRU	Q9wb02	SubName: Full=ORF2; Flags: Fragm...	<u>45</u>	0.002
>>>sptrembl:Q9WSW6_9VIRU	Q9wswo	SubName: Full=ORF2; Flags: Fragm...	<u>45</u>	0.002
>>>sptrembl:B3FWR6_9VIRU	B3fwr6	SubName: Full=ORF2; Flags: Fragm...	<u>38</u>	0.34

>>>>sptrembl:Q9WAZ2_9VIRU Q9waz2 SubName: Full=ORF2; Flags:

Fragmento ;. 2/2010

Longitud = 152

Puntuación = 49.3 bits (116), Esperado = 1e-04, Método: ajuste de matriz composicional.

Identidades = 23/23 (100%), Positivos = 23/23 (100%)

Query: 1 RVPEVSLHTAVKQFGLGTGRAM 23

RVPEVSLHTAVKQFGLGTGRAM

Sbjct: 14 RVPEVSLHTAVKQFGLGTGRAM 36

Q9WAZ2_9VIRU

Sin revisar ; 152 AA.

AC Q9WAZ2;

DT 01-NOV-1999, integrado en UniProtKB/TrEMBL.

DT 01-NOV-1999, versión de secuencia 1.

DT 09-FEB-2010, versión de entrada 23.

DE SubName: Full=ORF2;

ES 2 657 705 T3

```

DE   Flags: Fragmento;
OS   Torque teno virus.
OC   Virus; virus de ADNss ; Anelloviridae; Anelloviridae sin clasificar.
OX   NCBI_TaxID=68887;
RN   [1]
RP   SECUENCIA DE NUCLEÓTIDOS.
RC   TEJIDO=Sero;
RX   MEDLINE=99335592; PubMed=10405352; DOI=10.1006/viro.1999.9797;
RA   Hijikata M., Takahashi K., Mishiro S.;
RT   "Complete circular DNA genome of a TT virus variant (isolate name
RT   SANBAN) and 44 partial ORF2 sequences implicating a great degree of
RT   diversity beyond genotypes.";
RL   Virology 260:17-22(1999).
CC   -----
CC   Derechos de autor de UniProt Consortium, véase http://www.uniprot.org/terms
CC   Distribuido bajo la licencia de Creative Commons Attribution-NoDerivs
CC   -----
DR   EMBL; AB024353; BAA77420.1; -; Genomic DNA.
DR   InterPro; IPR004118; Gyrovir_VP2/TT_ORF2.
DR   Pfam; PF02957; TT_ORF2: 1.
PE   4: Previsto ;
FT   NON_TER      1      1
FT   NON_TER      152    152
SQ   SECUENCIA 152 AA; 16658 MW; 1DEC53175C043A17 CRC64;
      ARTPRRGVRA SRGRVPEVSL HTAVKQGFGL GTGRAMGKAL KKAMFLGRIY RKKRRRLPLSP
      LHSPPKARKL LRGMWRPPTQ NVSGQERSWY DSVFYSHAAF CGGDCVGHSL SYLATHLGRP
      PSAQPPPQLQ PPVIRRLPAL PAPPNPSGDR AA
  
```

zkb69.3.pep

nucleótidos 3-71
Longitud: 23

GAGGEFTHRS QGAIRARDWP GYG

```

BlastP2 de: zkb69.3.pep desde: 1 a: 23
comparado con la base de datos: uniprot ..
>>>sptrembl:Q9JGT3_9VIRU Q9jgt3 SubName: Full=PORF2a;. 12/2009 50 5e-05
>>>sptrembl:Q9JGS4_9VIRU Q9jgs4 SubName: Full=PORF2a;. 12/2009 50 5e-05
>>>sptrembl:Q98Y39_9VIRU Q98y39 SubName: Full=ORF2;. 2/2010 49 1e-04
>>>sptrembl:Q9WAY7_9VIRU Q9way7 SubName: Full=ORF2; Flags: Fragm... 49 1e-04
>>>sptrembl:Q9YR02_9VIRU Q9yr02 SubName: Full=Putative uncharact... 49 1e-04
>>>sptrembl:Q786D4_9VIRU Q786d4 SubName: Full=ORF2;. 2/2010 49 1e-04
>>>sptrembl:O70738_9VIRU O70738 SubName: Full=ORF2, ORF1 genes;... 49 1e-04
>>>sptrembl:Q9JG33_9VIRU Q9jg33 SubName: Full=Putative uncharact... 49 2e-04
>>>sptrembl:O90363_9VIRU O90363 SubName: Full=ORF2 protein;. 2/2010 49 2e-04
>>>sptrembl:Q9JGT0_9VIRU Q9jgt0 SubName: Full=PORF2a;. 12/2009 49 2e-04
>>>sptrembl:Q9YKL2_9VIRU Q9ykl2 SubName: Full=ORF2 protein;. 2/2010 49 2e-04
>>>sptrembl:Q9WFY6_9VIRU Q9wfy6 SubName: Full=ORF2;. 2/2010 49 2e-04
>>>sptrembl:Q9JGS7_9VIRU Q9jgs7 SubName: Full=PORF2a;. 12/2009 48 3e-04
>>>sptrembl:Q9DYC0_9VIRU Q9dyc0 SubName: Full=Putative uncharact... 48 3e-04
>>>sptrembl:Q9W7S4_9VIRU Q9w7s4 SubName: Full=Putative uncharact... 48 3e-04
>>>sptrembl:Q77S01_9VIRU Q77s01 SubName: Full=Putative uncharact... 48 3e-04
>>>sptrembl:B3FWR5_9VIRU B3fwr5 SubName: Full=ORF2; Flags: Fragm... 34 5.2
>>>sptrembl:B6SFP1_9VIRU B6sfp1 SubName: Full=ORF2; Flags: Fragm... 34 5.6

>>>>sptrembl:Q9JGT3_9VIRU Q9jgt3 SubName: Full=PORF2a;. 12/2009
Longitud = 49
Puntuación = 50.4 bits (119), Esperado = 5e-05, Método: ajuste de matriz composicional.
Identidades = 23/23 (100%), Positivos = 23/23 (100%)
Query: 1 GAGGEFTHRSQGAIRARDWPGYG 23
      GAGGEFTHRSQGAIRARDWPGYG
  
```

ES 2 657 705 T3

Sbjct: 24 GAGGEFTHRSQGAI RARDWPGYG 46

Q9JGT3_9VIRU Sin revisar ; 49 AA.
AC Q9JGT3;
DT 01-OCT-2000, integrado en UniProtKB/TrEMBL.
DT 01-OCT-2000, versión de secuencia 1.
DT 15-DEC-2009, versión de entrada 17.
DE SubName: Full=PORF2a;
GN Name=ORF2a;
OS Torque teno virus.
OC Virus; virus de ADNss ; Anelloviridae; Anelloviridae sin clasificar.
OX NCBI_TaxID=68887;
RN [1]
RP SECUENCIA DE NUCLEÓTIDOS.
RX MEDLINE=20417334; PubMed=10963344; DOI=10.1007/s007050070097;
RA Tanaka Y., Orito E., Ohno T., Nakano T., Hayashi K., Kato T.,
RA Mukaide M., Iida S., Mizokami M.;
RT "Identification of a novel 23kDa protein encoded by putative open
RT reading frame 2 of TT virus (TTV) genotype 1 different from the other
RT genotypes.";
RL Arch. Virol. 145:1385-1398(2000).
CC -----
CC Derechos de autor de UniProt Consortium, véase <http://www.uniprot.org/terms>
CC Distribuido bajo la licencia de Creative Commons Attribution-NoDerivs
CC -----
DR EMBL; AB030486; BAA90401.1; -; Genomic_DNA.
DR InterPro; IPR013267; TTV_ORF2a.
DR Pfam; PF08197; TT_ORF2a; 1.
PE 4: . Previsto ;
SQ SECUENCIA 49 AA; 5118 MW; 596E44680A5D863A CRC64;
MAEFSTPVRS EGATEGIPNV PRAGAGGEFT HRSQGAI RAR DWPGYGQGS

Figura 15

El ADN celular de L1236 se sometió a amplificación con templiPhi, seguida de de 2 rondas sucesivas de PCR con los cebadores jt34f 7 y 8

