

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 387 604**

51 Int. Cl.:
C12N 15/51 (2006.01)
C07K 14/08 (2006.01)
A61K 39/29 (2006.01)
G01N 33/576 (2006.01)
C12Q 1/70 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 96 Número de solicitud europea: **10005982 .3**
96 Fecha de presentación: **23.09.2002**
97 Número de publicación de la solicitud: **2343374**
97 Fecha de publicación de la solicitud: **13.07.2011**

54 Título: **Moléculas de ácidos nucleicos de HDV, sus fragmentos y aplicaciones**

30 Prioridad:
24.09.2001 FR 0112285

45 Fecha de publicación de la mención BOPI:
31.10.2012

45 Fecha de la publicación del folleto de la patente:
31.10.2012

73 Titular/es:
**ASSISTANCE PUBLIQUE - HÔPITAUX DE PARIS
(100.0%)
3 Avenue Victoria
75004 Paris, FR**

72 Inventor/es:
**DENY, PAUL;
RADJEF, NADJIA y
HUC-ANAIS, PATRICIA**

74 Agente/Representante:
PONS ARIÑO, Ángel

ES 2 387 604 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

Moléculas de ácidos nucleicos de HDV, sus fragmentos y aplicaciones

5 La presente invención se refiere a moléculas de ácidos nucleicos procedentes de nuevas cepas o aislamientos del virus de la hepatitis D que constituyen genotipos diferentes a los genotipos conocidos I, II y III, a sus fragmentos, a las proteínas correspondientes, así como a sus aplicaciones como reactivos de diagnóstico.

La presente invención se refiere igualmente a un procedimiento de diagnóstico sensible del virus de la hepatitis D (o virus de la hepatitis delta), así como a un método/procedimiento de seguimiento epidemiológico de las infecciones por HDV.

10 El virus de la hepatitis D (VHD o HDV, por "hepatitis D virus") o delta es un virus satélite de la hepatitis B. Este virus tiene una estructura particular: una estructura quimérica que asocia a componentes específicos del HDV (ARN vírico y proteínas de HD), una cubierta que comprende las tres glucoproteínas del HBV: grande (preS1-preS2-S), mediana (preS2-S) y pequeña (S). El diámetro medio de las partículas de HDV se sitúa entre el de las partículas de HBV maduras (partículas de DANE: 42 nm) y el de las cubiertas vacías de HBV (formas esféricas o filamentosas: 22 nm), y la densidad de flotación es de 1,24-1,25 g/cm³.

15 En el interior de los viriones, el ARN de HDV es circular y de polaridad negativa. El genoma más pequeño conocido de los virus que infectan a mamíferos, esta hebra monocatenaria circular cerrada tiene un porcentaje elevado de GC (60%).

El ARN de HDV se replica independientemente de HBV, cuyo papel se limita a proporcionar la cubierta del HDV. Las únicas proteínas encontradas (sHD y LHD) se codifican por el ARN antígenómico que, en la célula infectada, es completo, circular y pseudocatenario, sirve como intermedio de replicación y soporta la edición.

20 El ARN de HDV pertenece a un tipo de ribozima específico. La reacción de autoescisión requiere ARN y un catión divalente (Mg⁺⁺). La escisión crea un extremo 2',3'-fosfato cíclico y un extremo 5' hidroxilo.

Las ribozimas delta (genómica y antígenómica) tienen una configuración secundaria cercana a un pseudonudo. Las secuencias implicadas incluyen principal o exclusivamente secuencias situadas en 3' del sitio de autoescisión (aproximadamente 84 nucleótidos).

25 En el transcurso del ciclo vírico, el ARNm de HDV codifica una proteína que existe en dos formas: una proteína de 194-195 aminoácidos (forma "s" de pequeña) de 24 kilodalton (kDa) y una proteína de 214 aminoácidos (forma "L" de grande) de 27 kDa, que existen en proporciones variables. Estas proteínas portan la antigenicidad "delta" y se detectan en el hígado o el suero de pacientes o animales infectados (chimpancés, marmotas). Estas dos proteínas víricas sHD y LHD se inician en el primer ATG del marco de lectura abierto situado en la posición 1598 (según la numeración de Wang *et al.*, 1986 o 1987) del ARN antígenómico. En el transcurso de la replicación, aparece una mutación, dependiente de una enzima celular, la adenosina desaminasa dependiente de ARN bicatenarios, en la posición 1012, que transforma el codón de terminación ámbar (UAG) en el codón de triptófano (UGG), alargando en 3' el marco de lectura de 10 o 20 codones y confiriendo a las dos formas sHD y LHD propiedades diferentes.

30 El ARNm termina con una cola de poli(A), 15 nucleótidos después de la señal de consenso de poliadenilación AAUAAA (posiciones 954-959).

En el ciclo de replicación, las funciones de las proteínas de 24 y 27 kDa son opuestas: sHD activa la replicación vírica mientras que LHD la reprime y desempeña un papel en el ensamblaje de las partículas víricas. Esas proteínas están fosforiladas en los residuos de serina pero no glucosiladas (Tabla I). Están constituidas por dominios funcionales comunes y por un dominio específico de la proteína grande LHD.

40 Tabla I: Resumen y comparación de las funciones de las dos formas p24 y p27

Actividades bioquímicas y biológicas	p24(S)	p27(L)
Aminoácidos	195	214
Transactivación de la replicación	+	-
Transinhibición de la replicación	-	+
Dimerización y polimerización	+	+
Fijación a ARN	+	+
Estabilización de ARN	+	+
Localización nuclear	+	+
Ensamblado	-	+
Fosforilación	+	+(x6)
19 aa carboxiterminales específicos	-	+
Farnesilación	-	+

Brevemente, los diferentes dominios de estas dos proteínas son los siguientes:

* dominios comunes

5 - El dominio de polimerización, que comprende la secuencia entre los residuos aminoacídicos 13 y 48, formado por una disposición de leucina o isoleucina, organizado en hélice α de tipo "cremallera de leucina", implicado en la polimerización de proteínas, indispensable en (i) la transactivación de la replicación vírica por Ag-sHD, (ii) la inhibición de la replicación por Ag-LHD y (iii) el ensamblado de los complejos sHD-LHD en las cubiertas de HBV.

10 - La señal de localización nuclear (SLN), que implica dos secuencias de localización nuclear identificadas en la región 67-88, esenciales para translocar el Ag-sHD sintetizado en el citoplasma, y quizás la ribonucleoproteína después de su entrada en la célula, hacia el núcleo.

- el sitio de fijación a ARN que se basa en dos secuencias ricas en arginina localizadas entre los residuos 97 y 163, que permiten la unión de las proteínas sHD con el ARN genómico o antígenómico. Esta fijación es indispensable para que el Ag-sHD active la replicación.

* dominios específicos

15 Los 19-20 aminoácidos situados en el extremo COOH de la proteína grande tienen un papel importante en el ciclo del HDV. Efectivamente, estos aminoácidos (aa 195-214) intervienen en el ensamblado de las partículas víricas (Chang *et al.*, 1991). Esta actividad podría estar ligada en parte a la presencia de una cisteína en la posición 211 (Glenn *et al.*, 1992) conservada para todos los genomas víricos caracterizados hasta ahora. Esta cisteína, situada 4 aminoácidos antes del extremo COOH de la proteína, forma una secuencia "CXXX" y fija un agrupamiento farnesilo (Glenn *et al.*, 20 1992), cadena de 15 carbonos derivada del ácido mevalónico, mediante la acción de una farnesiltransferasa. Esta maduración postraduccional orienta las proteínas hacia las membranas celulares.

25 La proteína pequeña y grande se han diferenciado por otro lado mediante anticuerpos monoclonales (clon 9E4) (Hwang y Lai, 1993a). Estos anticuerpos solo reconocen a sHD (Lai *et al.*, 1993). Al estar incluida la secuencia aminoacídica de la proteína pequeña en la grande, estos resultados sugieren una conformación diferente entre sHD y LHD dentro de los 30 aminoácidos carboxiterminales de la proteína pequeña sHD, sugiriendo que el epítipo reconocido en sHD está enmascarado en LHD en condiciones no desnaturalizantes.

El HDV se transmite sobre todo por el contacto con agujas contaminadas y sangre, por tanto a través de los portadores de HDV o HBV.

30 En América del Norte y Europa Occidental, la hepatitis D se encuentra por tanto, sobre todo, en los usuarios de drogas intravenosas, hemofílicos y aquellos que han recibido transfusiones múltiples.

La epidemiología y los modos de contaminación se superponen en parte. Se estima globalmente en un 5% la proporción de portadores de Ag-HBs infectados por HDV. Sin embargo, se han comprobado disparidades de prevalencia geográfica y epidemiológica.

35 Existe una alta prevalencia de esta enfermedad en los portadores de la hepatitis B de ciertas regiones del mundo, que incluyen la cuenca amazónica de América del Sur, África central, el sur de Italia y los países de Oriente Medio.

40 En el portador mediterráneo, muy particularmente en el sur de Italia, Grecia y Oriente Medio, en donde la frecuencia de ser portador crónico del HBV es intermedia (1 a 5%), la infección por HDV es elevada. En estas regiones, se ha sugerido transmisión intrafamiliar, argumentada por estudios filogenéticos de los virus que infectan a los miembros de una misma familia (Niro *et al.*, 1999). En el sur de Italia, la prevalencia en el sujeto Ag-HBs positivo va disminuyendo, pasando de un 23% en 1987 a un 8% en 2000 (Gaeta *et al.*, 2000).

45 En África y en Asia, en donde la frecuencia de ser portador crónico de HBV es elevada (del 10 a 20%), así como en América del Sur y en las islas del Pacífico, donde es intermedia (1 a 5%), la distribución del HDV es paradójicamente dispar. En África, los estudios de seroprevalencia muestran una distribución muy heterogénea de los pacientes que tienen anticuerpos anti-HD, mientras que la prevalencia global de infección por HBV, estimada mediante la detección de Ag-HBs, se estabiliza entre el 12 y 14% (Roingeard *et al.*, 1992). Así, el ritmo variable de un 4% (zona norte de Senegal) a un 44% (afueras de Dakar) hace surgir los probables factores socioeconómicos implicados en la transmisión.

50 Los estudios de prevalencia del HDV han de interpretarse prudentemente. Efectivamente, en las poblaciones estudiadas, existe una inclusión preferida de pacientes aquejados de hepatopatías. En los pacientes aquejados de hepatitis aguda o crónica, la prevalencia de infección por HDV es mayor que en los portadores crónicos asintomáticos del HBV. Además, el estudio serológico de una infección por HDV se basa en la detección de Ag-HD y Ac anti-HD totales en el suero. Por ello, las infecciones agudas benignas, en el curso de las cuales se desarrollaría una producción transitoria aislada de IgM anti-HD, no se registrarían.

El HDV es responsable de hepatitis agudas y crónicas. Estas infecciones son particularmente graves y evolucionan más rápidamente hacia la cirrosis que las hepatitis B solas. Esta es una de las razones por las que es crucial un diagnóstico fiable de HDV asociado a HBV.

5 La infección por un HDV depende del HBV. Los aislamientos de HDV de regiones geográficamente diferentes muestran variabilidad genética. Actualmente, se han identificado tres genotipos y se han designado como genotipo I, II y III.

El genotipo sirve para determinar la epidemiología de las transmisiones víricas, permitiendo estudiar la distribución geográfica y podría correlacionarse con la potencia patogénica.

El HDV solo se desarrolla en los pacientes infectados también por HBV. Esta doble infección deriva o bien de una coinfección o bien de una sobreinfección:

10 - La coinfección está en el origen de la hepatitis aguda. El diagnóstico, realizado en el transcurso de una citolisis hepática, se basa en la detección de los marcadores de HDV asociados a la presencia de IgM anti-HBc. El Ag-HBs, generalmente presente, sería excepcionalmente negativo, justificando repetir las muestras para seguir la cinética de evolución de los marcadores. Es clásico comprobar la inhibición de la replicación de HBV por HDV. Las IgM anti-HBc atestiguan una infección reciente por HBV. El Ag-HD, muy precoz, se detecta raramente, teniendo en cuenta su fugacidad. Los anticuerpos aparecen de 2 a 3 semanas después del inicio de los síntomas: predominan las IgM anti-HD, pero mantienen una valoración moderada (<1:1000). Se observan dos picos de elevación de transaminasas, separados por dos a cinco semanas, en un 10 a 20% de las coinfecciones, que reflejan probablemente diferentes cinéticas de replicación de los virus. La coinfección se caracteriza por tanto por una gravedad de la hepatitis aguda a menudo superior a la ocasionada por el HBV aislado. Así, se describen hepatitis fulminantes en América del Sur y en el África subsahariana o en ciertas poblaciones. La evolución está generalmente marcada por una resolución de la hepatitis después de la fase aguda y, a semejanza del historial natural del HBV, solo un 5% de los pacientes coinfectados evolucionan hacia la cronicidad.

25 - La sobreinfección se caracteriza por la aparición de una seroconversión de HDV en un paciente portador crónico de Ag-HBs. La viremia de HDV precede a la aparición de anticuerpos anti-HDV en ausencia de detección de IgM anti-HBc. La detección de estos marcadores puede preceder a una elevación de las transaminasas en varios meses. En la fase aguda, la sobreinfección se traduce en una hepatitis fulminante en más de un 10% de los casos. Además, pasada la fase aguda, la sobreinfección conlleva frecuentemente (de 60 a 70%) una hepatitis crónica activa con evolución rápida hacia cirrosis. En la fase aguda de la sobreinfección, la detección de Ag-HD es seguida rápidamente por la aparición de anticuerpos que persisten a tasas elevadas. Al contrario que los modelos clásicos de infección vírica, se detectan simultáneamente IgG anti-HD e IgM anti-HD en las hepatitis B-delta crónicas.

30 La Tabla II siguiente resume la evolución de los marcadores B y delta en el transcurso de coinfecciones y sobreinfecciones

Coinfección por HDV	Fase aguda	Evolución	
		Crónica	Curación
Ag-HBs	+	+	-
IgM anti-HBc	+	-	-
Ag-HD	+ /-	-	-
IgM anti-HD	+/-	+	-
IgG anti-HD	+/-	+	+
ARN de HDV	+	+	-
Ag-HD intrahepático	+	+	-

Sobreinfección por HDV		Evolución	
		Crónica	Curación
Ag-HBs	+	+	-
IgM anti-HBc	-	-	-
Ag-HD	+ /-	-	-
IgM anti-HD	+	+	-
IgG anti-HD	+	+	+
ARN de HDV	+	+	-
Ag-HD intrahepático	+	+	-

La coinfección y sobreinfección son clínicamente indistintas. El diagnóstico virológico se basa habitualmente en los diferentes marcadores séricos. Más raramente, el Ag-HD puede detectarse en los cortes anatomopatológicos de la biopsia de hígado.

- 5 Los marcadores permiten seguir la evolución de la enfermedad hacia la curación o la cronicidad, decidir la puesta en tratamiento de un enfermo y evaluar la eficacia.

El HDV no se aísla en cultivo celular y el diagnóstico se basa por tanto esencialmente en la búsqueda de Ag-HD (ELISA, IF) o del genoma vírico (hibridación, PCR, PCR inmediata) por técnicas directas y por la detección de los anticuerpos de IgM anti-HD e IgG anti-HD por los métodos indirectos (ELISA).

- 10 - La búsqueda de Ag-HD intrahepático puede discutirse en hepatitis fulminantes, teniendo en cuenta la cinética de aparición de los marcadores séricos. Este examen tiene valor de referencia para estudiar la replicación de HDV, pero no puede usarse rutinariamente.
- El Ag-HD sérico se busca en suero en presencia de un agente disociante que expone el Ag-HD, incluido en la cubierta vírica que porta el Ag-HBs. La presencia de anticuerpo (Ac) anti-HD a niveles elevados (hepatitis crónicas) que fija los antígenos séricos perjudica la detección. Se han desarrollado técnicas de transferencia Western con fines de búsqueda. La presencia del virus en la sangre es transitoria y limitada a la fase precoz de la infección, y la posibilidad de detectar el Ag-HD disminuye los días posteriores a la aparición de los síntomas.
- 15
- Se usa la inmunocaptura para detectar las IgM anti-HD y la competición para las IgG anti-HD. Las técnicas ELISA han usado en primer lugar como antígeno Ag-HD del suero o el hígado de pacientes o animales infectados. Las nuevas pruebas se basan en Ag-HD recombinantes o péptidos sintéticos.
- 20
- Las técnicas de hibridación o PCR-TI permiten la detección de ARN genómico después de la extracción de ácidos nucleicos y la desnaturalización de las estructuras secundarias. Se han descrito varios sistemas de cebadores: su elección es determinante, pues la variabilidad genética en las regiones denominadas conservadas puede conducir a falsos negativos si los cebadores elegidos no son apropiados para las cepas víricas en circulación. La elección de los cebadores de PCR debe tener en cuenta la epidemiología local de los genotipos descritos y es primordial conocer bien la distribución de estos genotipos en el mundo.
- 25

No obstante, tanto en caso de coinfección como en caso de sobreinfección, el Ag-HD es de hecho difícil de detectar, mientras que la viremia precede a la aparición de anticuerpos.

- 30 En este contexto, y especialmente a causa del descubrimiento de nuevos genotipos, son necesarios reactivos nucleicos y proteicos que permitan el diagnóstico de HDV cualquiera que sea el genotipo.

Efectivamente, el estudio de las secuencias nucleotídicas de HDV por diferentes equipos en el mundo solo ha permitido diferenciar, hasta ahora, tres genotipos distintos:

- El genotipo 1, que es el más frecuente y el más ampliamente extendido en el mundo. Desde su descripción inicial (chimpancé infectado experimentalmente) por Wang (Wang *et al.*, 1986; Wang *et al.*, 1987), varios grupos han secuenciado el genoma de HDV de aislamientos geográficos diferentes. La primera secuencia de HDV en el hombre se describió en 1987, en los Estados Unidos, por S. Makino *et al.*, en un paciente toxicómano (Makino *et al.*, 1987). El genotipo I está ampliamente extendido en Italia, Estados Unidos, Taiwán, Nauru, Francia, Líbano, China (Makino *et al.*, 1987; Chao *et al.*, 1991 b; Imazeki *et al.*, 1991; Lee *et al.*, 1992 ; Niro *et al.*, 1997 y Shakil *et al.*, 1997). Dentro del genotipo I, se describe un porcentaje de similitud nucleotídica de más de un 85%.
- 35

- 5 - Un aislamiento japonés (Imazeki *et al.*, 1990; Imazeki *et al.*, 1991) es el prototipo de un 2ª subgrupo de HDV. Este genotipo II, que en un primer momento se había descrito solo en Japón y Taiwán (Imazeki *et al.*, citados anteriormente; Lee *et al.*, 1996b) parece tener una distribución geográfica mucho más grande. Particularmente, se han caracterizado igualmente secuencias de genotipo I y de genotipo II procedentes de Yakutia (Rusia). Por último, algunos autores se apoyan en una diversidad intragenotípica para permitir la división en los subtipos IIA (Imazeki *et al.*, 1990; Imazeki *et al.*, 1991; Lee *et al.*, 1996b), IIB (Wu *et al.*, 1998; Sakugawa *et al.*, 1999) y IIC. En ciertos países, la infección por virus de genotipo II se asociaría a hepatitis menos graves que las causadas por los HDV de genotipo I o III (Wu *et al.*, 1995b).
- 10 - En 1993, se describió un 3º grupo para los genomas de virus peruanos y colombianos (Casey *et al.*, 1993a). El genotipo III solo se ha descrito en América del Sur, y más particularmente en la cuenca amazónica, asociado a hepatitis graves, hasta epidemias de hepatitis fulminantes con esteatosis microvesicular (Casey *et al.*, 1993a; Casey *et al.*, 1996b) y a una morbilidad y mortalidad elevadas. En esta región geográfica, se observa que el HDV de genotipo III está asociado de manera privilegiada a HBV de genotipo F. Se han aislado recientemente otros aislamientos de este grupo en Venezuela (Nakano *et al.*, 2001).
- 15 Comparando el conjunto de genomas, se describen de 2 a 4 regiones conservadas (Chao *et al.*, 1991b). Dos se encuentran siempre y están centradas alrededor de los sitios de autoescisión de los genomas y antígenomas implicados en las actividades autocatalíticas. Las otras dos regiones conservadas están situadas en el marco de lectura que codifica la proteína HD (Chao *et al.*, 1991b).
- 20 No obstante, las técnicas de detección dependen de la variabilidad genética del virus buscado; no permitiendo los reactivos conocidos, especialmente a partir de las secuencias específicas de genotipo I, II o III, detectar infecciones de HDV variante y especialmente los HDV de genotipo diferente de los anteriormente citados.
- En consecuencia, las técnicas de detección precisadas anteriormente tienen el riesgo de desembocar en resultados negativos tanto a nivel nucleico como en lo referente a la respuesta de anticuerpos.
- 25 El descubrimiento y la consideración de nuevas variantes son importantes para poner a punto reactivos de detección y de diagnóstico de hepatitis D (serodiagnóstico, PCR, hibridación) suficientemente sensibles y específicos, es decir, que no conduzcan a resultados falsos negativos o falsos positivos: efectivamente, puede observarse actualmente una disociación IgM anti-HD positivo / ARN-HDV negativo en el contexto de una hepatopatía grave.
- 30 En el marco de sus trabajos, los inventores han caracterizado ahora, sorprendentemente, que la diversidad genética del HDV era significativamente mayor que la anteriormente descrita, lo que tiene consecuencias sobre la fiabilidad del diagnóstico.
- Han descubierto particularmente 9 nuevas secuencias completas de HDV (tres originarias de Yakutia y seis originarias de África) que circulan igualmente en Île de France y que no pertenecen a ninguno de los genotipos conocidos.
- El análisis de estos nuevos aislamientos:
- confirma la existencia de una variabilidad mucho mayor del HDV que la descrita hasta ahora,
 - 35 - pone en cuestión la clasificación de los HDV en solo tres genotipos,
 - ha conducido a los inventores a proponer un algoritmo de PCR-RFLP a partir de una región parcial del genoma para el genotipado de HDV y
 - ha conducido a los inventores a poner a punto reactivos adaptados a un diagnóstico fiable de infecciones por HDV cualquiera que sea el genotipo, mientras que anteriormente se observaban numerosos resultados falsos negativos
 - 40 (existencia de nuevos genotipos).
- Los inventores tenían por tanto como objetivo proporcionar moléculas de ácidos nucleicos de HDV capaces de permitir la detección de una variante HDV con relación a los tres genotipos anteriormente descritos.
- La presente invención tiene por tanto como objeto moléculas de ácido nucleico aisladas, caracterizadas porque se seleccionan del grupo constituido por:
- 45 - el genoma completo de la variante de HDV denominada dFr48, que presenta la secuencia SEQ ID NO: 21 y
 - el genoma de un HDV que presenta una divergencia o distancia genética $\leq 15\%$ con la secuencia SEQ ID NO: 21.
- Según un modo de realización ventajoso de dichas moléculas, la región R0 se obtiene preferiblemente mediante amplificación del ARN de HDV con los cebadores 900S (SEQ ID NO:33) y 1280AS (SEQ ID NO:34).
- 50 En el sentido de la presente invención, se entiende por molécula de ácido nucleico una molécula de ADNc o ARN que presenta una de las secuencias genómicas de HDV tales como las que se definen anteriormente y sus secuencias complementarias codificantes y anticodificantes.

La presente invención tiene igualmente como objeto moléculas de ácido nucleico aisladas que comprenden al menos uno de los fragmentos de secuencias de una variante HDV tal como se definen anteriormente, seleccionados del grupo constituido por:

- a) el fragmento R0 del HDV variante aislado siguiente: dFr48, que presenta la secuencia siguiente: SEQ ID NO:50,
- 5 b) el fragmento R1, que se extiende de las posiciones 307 a 1289 del genoma de HDV,
- c) el fragmento R2, que se extiende de las posiciones 889 a 328 del genoma de HDV,
- d) el fragmento R3, que se extiende de las posiciones 1486 a 452 del genoma de HDV,
- e) el fragmento R'1, que se extiende de las posiciones 305 a 1161 del genoma de HDV,
- f) el fragmento R'2, que se extiende de las posiciones 984 a 331 del genoma de HDV,
- 10 g) el fragmento R644, que se extiende de las posiciones 889 a 446 del genoma de HDV,
- h) el fragmento G910, que se extiende de las posiciones 1206 a 929 del genoma de HDV,
- i) el fragmento p910, que se extiende de las posiciones 553 a 1550 del genoma de HDV,
- j) el ADNc que codifica la proteína sHD, de secuencia SEQ ID NO:24,
- k) el ADNc que codifica la proteína LHD, de secuencia SEQ ID NO:22.
- 15 En el sentido de la presente invención, las posiciones de los fragmentos en el genoma de HDV se indican en el genoma circular en orientación genómica, según la numeración de Wang *et al.*, 1986 o 1987.

La invención engloba igualmente fragmentos nucleotídicos complementarios de los precedentes, así como fragmentos modificados con relación a los precedentes mediante la retirada o adición de nucleótidos en una proporción de aproximadamente un 15%, con relación a la longitud de los fragmentos anteriores, y/o modificados a nivel de la naturaleza de los nucleótidos a condición de que los fragmentos nucleotídicos modificados conserven la capacidad de hibridación con las secuencias de ARN genómicas o antigenómicas de los aislamientos tales como se definen anteriormente.

Efectivamente, estas diferentes cepas víricas, en un mismo paciente en un momento dado, muestran una población heterogénea de moléculas de ARN de HDV; además, en el transcurso de una infección crónica, además de las heterogeneidades observadas a nivel del sitio de edición (posición 1012), son susceptibles a la aparición mutaciones. Las secuencias víricas parecen evolucionar dentro de las poblaciones víricas con una tasa de sustitución variable de 3×10^{-2} a 3×10^{-3} por nucleótido y por año.

Ciertos de estos fragmentos son específicos y se usan como sonda o como cebador; hibridan específicamente con una cepa de una variante de HDV, tal como se define anteriormente, o con una cepa emparentada; se entiende por HDV emparentado una variante tal como se define anteriormente, un HDV que presente una divergencia genética $\leq 15\%$.

Dichos fragmentos se usan para la detección y el seguimiento epidemiológico de infecciones por HDV. Por ejemplo, el fragmento R0 se usa para la detección (por PCR-TI) y el genotipado (por PCR-RFLP) de HDV. Los demás fragmentos que cubren la totalidad del genoma de HDV se usan para la caracterización molecular de los HDV variantes; el análisis filogenético de la secuencia completa del genoma o de fragmentos de éste, correspondientes particularmente a R0 o R2 permite vincular los perfiles observados en PCR-RFLP con un genotipo dado o caracterizar nuevos genotipos.

En consecuencia, la presente invención tiene igualmente como objeto un método de detección de una variante HDV según la invención, mediante hibridación y/o amplificación, realizada a partir de una muestra biológica, estando caracterizado dicho método porque comprende:

- 40 (1) una etapa de extracción del ácido nucleico a detectar, que pertenece al genoma del virus eventualmente presente en la muestra biológica,
- (2) la realización de al menos una amplificación génica con la ayuda de un par de cebadores seleccionados del grupo constituido por los cebadores capaces de amplificar una de las regiones siguientes del ARN genómico de HDV: R0, R1, R2, R3, R644, G910, p910, R'1 y R'2 y
- 45 (3) el análisis del producto amplificado por comparación con una de las moléculas de secuencia SEQ ID NO:21 correspondiente al genoma completo de la variante denominada dFr48.

Ventajosamente, la etapa (3) del análisis puede ejecutarse mediante restricción, secuenciación o hibridación; en este último caso, la sonda usada (especialmente en los chips de ADN) será ventajosamente un fragmento de 15 a 20 nucleótidos específico de dichos fragmentos amplificados.

ES 2 387 604 T3

Según un modo de ejecución ventajoso de dicho método, los cebadores específicos para la amplificación de las regiones R0, R1, R2, R3, R644, G910, p910, R'1 y R'2, ejecutada en la etapa (2), se seleccionan del grupo constituido por:

- 5 - los cebadores 900S (SEQ ID NO:33) y 1280AS (SEQ ID NO:34) para la amplificación de R0 (aproximadamente 400 pb),
- los cebadores 320S (SEQ ID NO:39) y 1280AS (SEQ ID NO:34) para la amplificación del fragmento R1 (aproximadamente 960 pb),
- los cebadores 900S (SEQ ID NO:33) y 320AS (SEQ ID NO:45) para la amplificación de R2 (aproximadamente 1100 pb), que contiene el gen sHD correspondiente a las posiciones 1598-950,
- 10 - los cebadores 1480S (SEQ ID NO:46) y 440AS (SEQ ID NO:47) para la amplificación de R3 (aproximadamente 650 pb),
- los cebadores 900S (SEQ ID NO:33) y 420AS (SEQ ID NO:40) para la amplificación de la región R644 (aproximadamente 1250 pb) del aislamiento dFr644,
- 15 - los cebadores 318S (SEQ ID NO:35) y 1150AS (SEQ ID NO:36) para la amplificación de R'1 (aproximadamente 850 pb),
- los cebadores 960S (SEQ ID NO:37) y 345AS (SEQ ID NO:38) para la amplificación de R'2 (aproximadamente 1050 pb),
- los cebadores R910S (SEQ ID NO:41) y R910As (SEQ ID NO:42) para la amplificación de la región G910 (aproximadamente 1400 pb) del aislamiento dFr910,
- 20 - los cebadores S1910R (SEQ ID NO:43) y As1910R (SEQ ID NO:44) para la amplificación de la región p910 (aproximadamente 650 pb) del aislamiento dFr910.

La presente invención tiene igualmente como objeto un método de detección y genotipado de HDV a partir de una muestra biológica, estando caracterizado dicho método porque comprende:

- (a) una etapa de extracción del ácido nucleico que pertenece al genoma del virus HDV,
- 25 (b) una etapa de amplificación de la región R0 delimitada por las posiciones 889 a 1289 del genoma de HDV,
- (c) un primer tratamiento de las moléculas de ácidos nucleicos amplificadas con las enzimas de restricción SmaI y XhoI, produciendo un primer conjunto de fragmentos de restricción, y
- (d) un segundo tratamiento de las moléculas de ácidos nucleicos por la enzima de restricción SacII, produciendo un segundo conjunto de fragmentos de restricción,
- 30 (e) el análisis combinado de los dos conjuntos de fragmentos de restricción producidos por RFLP (polimorfismo de longitud de fragmento de restricción), de manera que se detecte la presencia y/o se determine el tipo de HDV presente en dicha muestra biológica.

Según un modo de ejecución ventajoso de dicho método, la etapa (b) de amplificación se ejecuta ventajosamente con los cebadores 900S (SEQ ID NO: 33) y 1280AS (SEQ ID NO:34).

- 35 El método según la invención permite definir nuevos perfiles de restricción y clasificar los HDV en siete genotipos distintos.

Según otro modo de ejecución ventajoso de dicho método, comprende además:

- (f) la amplificación de moléculas de ácido nucleico de dicha muestra por PCR-TI con los cebadores 900S (SEQ ID NO:33) y 320AS (SEQ ID NO: 45) de manera que se amplifique la región R2 y
- 40 (g) la secuenciación directa de la región R2 amplificada y la comparación con una de las moléculas de ARN de secuencias SEQ ID NO: 1, 6, 11, 16, 21 y 26, correspondientes respectivamente a los genomas completos de los aislamientos o variantes denominados respectivamente dFr45, dFr47, dFr73, dFr910, dFr48 y dFr644, por ejemplo mediante análisis filogenético.

- 45 Cuando se observan perfiles inhabituales, esta etapa complementaria permite caracterizar genotipos nuevos. Efectivamente, estos análisis complementarios de PCR-RFLP permiten vincular el nuevo perfil observado con un genotipo dado, o caracterizar un nuevo genotipo, mediante análisis filogenético.

La presente invención tiene igualmente como objeto un vector recombinante, especialmente un plásmido, que comprende un inserto constituido por una molécula de ácido nucleico tal como se define anteriormente.

La presente invención tiene igualmente como objeto una célula transformada por una molécula de ácido nucleico tal como se define anteriormente.

5 La presente invención tiene igualmente como objeto los productos de traducción codificados por una de las moléculas de ARN de secuencia SEQ ID NO: 21 correspondiente al ARN genómico completo de la variante denominada dFr48 o por su secuencia complementaria codificante o anticodificante.

La presente invención tiene igualmente como objeto las proteínas codificadas por el genoma de una variante HDV tal como se define anteriormente.

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

- la proteína LHD de dFr48, que presenta respectivamente la secuencia SEQ ID NO:23 y
- 10 - la proteína sHD de dFr48, que presenta la secuencia SEQ ID NO:25.

La presente invención tiene igualmente como objeto un péptido caracterizado porque está constituido por un fragmento de una proteína tal como se define anteriormente, seleccionado del grupo constituido por:

- el péptido A constituido por los 19 aminoácidos del extremo carboxiterminal de la secuencia SEQ ID NO:23,
- el péptido C constituido por los 9 aminoácidos precedentes de la secuencia SEQ ID NO:59 (SEQ ID NO:64).
- 15 Dichos péptidos son útiles para el diagnóstico indirecto (serología) de una infección por HDV, especialmente mediante un procedimiento inmunoenzimático (ELISA):
- el péptido B, que está conservado, permite la detección del conjunto de variantes según la invención y del genotipo II de HDV, y
- el péptido C es específico de las diferentes variantes de HDV según la invención.

20 La presente invención tiene además como objeto el uso de una molécula de ácido nucleico tal como se define anteriormente o de una proteína tal como se define anteriormente para la preparación de un kit de detección y genotipado de HDV.

Además de las disposiciones precedentes, la invención comprende también otras disposiciones que surgirán de la descripción siguiente, que se refiere a ejemplos de ejecución de la presente invención, así como a los dibujos anexos en los que:

- La figura 1 representa el árbol filogenético de la región R0 obtenido mediante el procedimiento del vecino más cercano "neighbor-joining". Los números en cursiva indican los valores de aleatorización muestral (bootstrapping) (BV) sobre 10^4 remuestras y el signo ▲ indica los BV < 50%. La escala representa el número de sustituciones de nucleótidos por sitio.
- 30 - La figura 2 representa el árbol filogenético de las regiones R0 de HDV obtenido mediante el procedimiento de máxima parsimonia. Los aislamientos Perú1, Perú2 y Colombia se eligen como "grupo externo". Los números en cursiva indican los valores de aleatorización muestral (BV) sobre 10^4 remuestras,
- La figura 3 ilustra los datos clínicos de cada uno de los 6 pacientes infectados por los aislamientos de HDV de origen africano. * indica respectivamente las PCR 6S/6As y PCR R0,
- 35 - La figura 4 representa el árbol filogenético de los genomas completos de HDV obtenido mediante el procedimiento del vecino más cercano "neighbor-joining". Los números en cursiva, al nivel de cada nodo, indican los valores de aleatorización muestral (BV) sobre 10^4 remuestras. La escala representa el número de sustitución de nucleótidos por sitio.
- La figura 5 representa el árbol filogenético de los genomas completos de HDV, obtenido mediante el procedimiento de máxima parsimonia. Los números en cursiva, al nivel de cada nodo, indican los valores de aleatorización muestral (BV) sobre 10^4 remuestras.
- 40 - La figura 6 representa el alineamiento de secuencias de aminoácidos de proteínas delta de los 6 aislamientos de origen africano (líneas 7, 8, 9, 10, 11 y 12) con las secuencias conocidas de genotipo I (líneas 13, 14 et 15), genotipo II (líneas 3, 4, 5 y 6), genotipo III (línea 16) y TW2b/Miyako (líneas 1 y 2), con la ayuda del software Clustal (versión 1.8). El aminoácido en posición 196 de la proteína p27 corresponde al codón de terminación de la proteína p24 (Z) o al codón de triptófano (W) que conduce a la síntesis de la proteína p27 que se extiende de los aminoácidos 1 a 215.
- 45

No obstante, debe entenderse que estos ejemplos se dan únicamente a modo de ilustración del objeto de la invención, del que no constituyen en manera alguna una limitación.

EJEMPLO 1: MATERIALES Y MÉTODOS**1- Pacientes y muestras**

Se analizaron 22 sueros procedentes de sujetos seguidos en diferentes centros hospitalarios de la región parisina. Los pacientes eran portadores crónicos de Ag-HBs. El diagnóstico de infección delta se efectuó mediante la búsqueda de los marcadores serológicos (Ag-HD, anti-AgHD de tipo IgM e IgG) y la detección del genoma vírico de HDV por PCR-TI. No se detectó Ag-HD en ninguno de los sueros analizados. Los anticuerpos anti-AgHD de tipo IgM atestiguan la cronicidad de la infección delta, y los de tipo IgG se encontraron en todos los pacientes. Se caracterizó la totalidad del genoma de HDV en 6 pacientes. Todos los sueros se conservaron a -80°C hasta la extracción del ARN vírico.

2- Extracción de ARN de HDV

Para extraer el ARN de HDV, se añade un volumen de 250 µl de suero a 75 µl de reactivo Trizol LS (Gibco BRL, Life Technologies). Después de 30 segundos de homogeneización, se incuba la mezcla durante 5 min a temperatura ambiente. La extracción de lípidos se efectúa añadiendo 200 µl de cloroformo enfriado a +4°C. Después de una nueva homogeneización con vórtex, se incuban los tubos y entonces se centrifugan a 14.000 revoluciones por minuto (rpm) durante 10 min a +4°C. Se transfiere la fase acuosa a tubos de extracción y se precipitan los ARN con 500 µl de isopropanol frío en presencia de 1 µg de glucógeno. Después de 15 min de homogeneización, se centrifugan las muestras a 14.000 rpm durante 10 min a +4°C. Después de aclarar con etanol al 70%, se centrifugan de nuevo los tubos a 14.000 rpm durante 10 min a +4°C. Se secan los sedimentos en campana a temperatura ambiente y se recogen entonces en 100 µl de agua estéril que comprenden un inhibidor de ribonucleasas (RNasin, Promega). En esta etapa, se toman precauciones para evitar eventuales contaminaciones de tampones y muestras por ribonucleasas.

3- Síntesis de ADN complementario (ADNc)

Esta etapa consiste en sintetizar una hebra de ADN complementario del ARN de HDV mediante transcripción inversa.

Para eliminar la estructura secundaria del ARN de HDV, se añaden 5 µl de ARN extraído anteriormente a una mezcla de reacción que contiene 5 µl (o 0,5 pmoles) de trifosfatos de desoxinucleótidos (dNTP) y 1 µl (0,4 pmoles) de hexanucleótidos aleatorios. Se desnaturalizan a continuación los ARN durante 3 min a 95°C. Para fijar los ARN desnaturalizados, se congelan los tubos inmediatamente en etanol enfriado a -20°C. Se añaden al ARN desnaturalizado 10 µl de una mezcla de reacción que contiene 2,5 µl de ditiotretitol (DTT), 100 unidades (U) de transcriptasa inversa Superscript II, (Gibco BRL, Life Technologies) y su tampón de reacción, así como 20 U de inhibidor de ribonucleasas (RNasin, Promega). La reacción de transcripción inversa se efectúa a 42°C durante 45 min, entonces se detiene mediante incubación a 94°C durante 5 min. Se conservan a continuación los ADNc a -80°C.

4- Amplificación génica

Se efectúa la amplificación de ADNc de manera exponencial por PCR (reacción en cadena de la polimerasa). Se usaron dos tipos de polimerasas: la polimerasa AmpliTaq Gold (*Thermophilus aquaticus*) (PE Applied Biosystems) y la polimerasa Pwo (*Pyrococcus woesei*) o el sistema de PCR Expand™ High Fidelity (Roche).

Se efectuó la amplificación a partir de 5 µl de ADNc, que se añaden a una mezcla de reacción de PCR que contiene: 0,25 pmoles/µl de cebador codificante y anticodificante (Tabla III), 200 µmoles de cada dNTP, MgCl₂ 1,5 mM, 1 U de AmpliTaq Gold o 2,6 U de polimerasas Expand™ en presencia de los tampones de PCR correspondientes. Se realizó la reacción de PCR en un termociclador (PCR Sprint, Hybaid, Coger) según el protocolo siguiente: desnaturalización de los híbridos de ADNc-ARN a 94°C durante 9 min, seguido de 40 ciclos sucesivos, que comprende cada uno una desnaturalización de los ADNc a 94°C durante 45 s, una hibridación de los cebadores (900S/1280As o 6S/6As) a 58°C durante 30 s y la síntesis de la hebra complementaria gracias a la polimerasa mediante elongación a 72°C durante 45 s. Por último, una elongación final a 72°C durante 4 min y 30 s a 72°C.

Tabla III: Secuencias de los cebadores usados para las reacciones de PCR y su posición en el genoma de HDV

Cebadores	Secuencia 5'→3'	Número de identificación	Posiciones*
6S	gaggaaagaaggacgcgagacgcaa	SEQ ID NO:31	904-929
6As	accccctgaaggtggatcga	SEQ ID NO: 32	1141-1121
900S	catgccgaccogaagaggaaag	SEQ ID NO: 33	889-911
1280As	gaaggaaggccctcgagaacaaga	SEQ ID NO:34	1289-1265
318S	ctccagaggaccccttcagcgaac	SEQ ID NO:35	305-328
1150As	cccgcgggtggggatgtgaacc	SEQ ID NO:36	1161-1138
960S	gtacactcgaggagtgaaggcg	SEQ ID NO:37	962-984
345As	tctgtcgtgaagggctct	SEQ ID NO:38	331-311
320S	ccagaggaccccttcagcgaac	SEQ ID NO 39	307-328
420As	aacaccctcctgtagcccc	SEQ ID NO: 40	446-427
R910S	ccggagtctctctctctctcc	SEQ ID NO:41	1206-1227
R910AS	gttcgcgtcegagtctctctctt	SEQ ID NO:42	929-907
S1910R	gagcttcttcgattcggac	SEQ ID NO:43	1531-1550
AS1910R	gactggtcccctcatgttc	SEQ ID NO:44	572-553

Según la numeración de Wang *et al.* (Nature, 1986, 323, 508-514; Nature, 1987, 328, 456)

4.1 – Estrategia de amplificación del genoma vírico de HDV

- 5 El par de cebadores 6S y 6As permite amplificar un fragmento de ADN correspondiente al extremo carboxiterminal del gen que codifica el antígeno delta.
- Se amplificó la región R0, que comprende el extremo carboxiterminal del gen que codifica Ag-HD y una parte de la región no codificante, para todas las muestras con la ayuda de los cebadores 900S (SEQ ID NO:33) y 1280As (SEQ ID NO:34). En el cebador 900S usado se eliminaron 7 nucleótidos del extremo 5' con relación al usado por Casey *et al.*, 1993a, anteriormente citado, para la clasificación de los genotipos de HDV.
- 10 La selección de estos cebadores permite, sorprendentemente, la amplificación de un fragmento capaz de permitir distinguir los genotipos conocidos (I, II y III) de los genotipos nuevos.
- Se obtuvieron las secuencias completas del genoma vírico de HDV de cuatro muestras (dFr45, dFr47, dFr48 y dFr73) mediante amplificación de dos regiones superpuestas R'1 (850 bases) y R'2 (1050 bases), respectivamente, con la ayuda de los pares de cebadores 318S/1150AS y 960S/345AS.
- 15 Para la muestra dFr644, la variabilidad observada en la región correspondiente a los cebadores anteriormente descritos condujo a amplificar la región 644 (R644) con la ayuda de un par de cebadores específicos: 900S y 480AS.
- 20 Para la muestra dFr910, la secuencia nucleotídica R0 permitió definir cebadores nuevos y específicos de la muestra para amplificar el genoma completo. Se eligieron dos pares de cebadores: los cebadores R910S y R910AS, que amplifican un fragmento de 1400 bases correspondiente a la región G910. Fue indispensable otro par de cebadores S 1910R y AS 1910R, que amplifican un fragmento de 650 bases (región p910), para cubrir la totalidad del genoma.
- Se efectuó la amplificación de las diferentes regiones (R1, R2, R3, R644, R'1, R'2, G910 y p910) como se describe anteriormente. Las temperaturas de hibridación y elongación, así como el tiempo de elongación usado para cada una de las PCR, se indican en la Tabla IV.

Tabla IV: Amplificación de los diferentes fragmentos del genoma

Regiones amplificadas	Tamaño de los fragmentos (bases)	Temperatura de hibridación (°C)	de Temperatura de elongación (°C)	de Tiempo de elongación
RI	960	62	72	1 min 15 s
R2	1100	56	72	1 min 30 s
R3	650	50	72	1 min
R644	1250	58	72	40 s
G910	1400	58	72	1 min 40 s
p910	650	58	72	1 min
R'1	850	63	72	1 min
R'2	1050	60	72	1 min 20 s

5-Análisis de los productos de amplificación

- 5 Se deposita un volumen de 8 µl del producto de PCR en presencia de 2 µl de solución de deposición en un gel al 1,3% de agarosa preparado en tampón Tris-borato/EDTA 0,5X y que contiene bromuro de etidio (BET) 0,5 µg/ml. Se efectúa la electroforesis en tampón TBE 0,5X. Se efectúa la migración en presencia de un marcador de tamaño (Raoul™, Appligène). El fragmento amplificado se visualiza bajo rayos ultravioleta a 312 nm y se fotografía.

6- Clonación y secuenciación de genomas de HDV

- 10 Antes de la etapa de clonación y secuenciación, se purificaron los productos de amplificación con el fin de eliminar todas las trazas de sales y enzimas.

6.1- Elongación por polimerasa Taq estándar

Se efectúa esta etapa en el caso en que la amplificación del producto se haya realizado con polimerasa Pwo. Permite añadir residuos de desoxiadenosina (A) a los extremos 3' de los productos de PCR, debido a que la polimerasa Pwo, dotada de actividad exonucleasa 5' → 3', disminuye la incorporación de desoxiadenosinas.

- 15 Se añade un volumen de 10 µl de ADN purificado a una mezcla de reacción de 70 µl que contiene: dNTP 0,2 mM, MgCl₂ 1,5 mM, tampón 1X y 2,5 U de polimerasa Taq (Perkin Elmer). Se efectuó la elongación a 72°C durante 30 minutos. Los productos de PCR experimentan entonces una nueva purificación con fenol-cloroformo y una precipitación con etanol, y se recogen entonces con 10 µl de agua estéril.

6.2- Clonación en el vector pCRII-TA-cloning (Invitrogen)

- 20 Se usa la clonación para confirmar la secuencia nucleotídica del ADN amplificado. Se realiza usando el vector pCRII (Invitrogen).

- 25 El vector pCRII se presenta en forma lineal. Posee residuos de desoxitimidina (T) que permiten la clonación del producto amplificado gracias a los residuos complementarios de desoxiadenosina (A) añadidos por la polimerasa Taq. Posee igualmente las secuencias promotoras Sp6 y T7, dos sitios de restricción EcoRI que rodean al sitio de inserción del producto de PCR y los genes de resistencia a ampicilina y kanamicina. Una fracción del gen lacZα, que codifica la β-galactosidasa, facilita la selección de los recombinantes por el color de las colonias. Efectivamente, los plásmidos que han integrado el inserto no expresan el gen lacZα. Las colonias bacterianas son entonces blancas en presencia de sustrato de β-galactosidasa (X-Gal o 5-bromo 4-cloro-3-indolil-β-galactoside, Roche) y un inductor del gen (IPTG o isopropiltio-β-D-galactósido, Roche). Así, las bacterias recombinantes se seleccionan gracias a su resistencia a la ampicilina y a un cribado azul-blanco.

La relación molecular elegida de inserto/vector es de 3/1 y el volumen de producto de PCR usado es variable, según la cantidad de ADN estimada por electroforesis en gel de agarosa como se describe anteriormente. La mezcla de reacción de 10 µl contiene 50 ng de vector pCRII, la cantidad de inserto correspondiente, 4 U de ADN ligasa T4 y el tampón de ligasa 1x. La reacción de ligamiento se realiza durante 18 horas a 14°C. Se conservan a continuación los tubos a +4°C.

- 35 Se conservan las bacterias *Escherichia coli* TOP10F' (Invitrogen), convertidas en competentes por tratamiento con cloruro de calcio, a -80°C listas para su uso. Se pone un contacto un volumen de 50 µl de bacterias competentes con 3 µl de la disolución de ligamiento durante 30 minutos en hielo. Un choque térmico (30 s a 42°C) hace penetrar el ADN plasmídico en las bacterias, que se devuelven inmediatamente al hielo varios minutos antes de incubarse durante 1 hora a 37°C en 250 µl de medio SOC (triptona al 2%; NaCl 10 mM; KCl 2,5 mM; MgCl₂ 10 mM; glucosa 20 mM, extractos de

levadura 5 g/l). Se aíslan a continuación las colonias en placas Petri que contienen agar LB (medio de Luria-Bertani), suplementado con ampicilina (50 µl/ml), y se distribuyen 40 µl de X-Gal (40 mg/ml) y 40 µl de IPTG (100 mM).

6.3-Extracción plasmídica y análisis del inserto

5 Se siembran las colonias blancas en caldo de LB-ampicilina (50 µl/ml) y se incuban durante 18 horas a 37°C con agitación. Se mantiene una colonia azul, es decir que no ha insertado el fragmento, como control negativo de ligamiento.

10 Se realiza la extracción plasmídica con la ayuda de un kit comercial QIAprep® Spin Miniprep (Qiagen). Brevemente, después de la centrifugación (3000 rpm a +4°C) y eliminación del sobrenadante, se suspende el sedimento bacteriano en 250 µl de tampón (Tris-HCl 50 mM pH 8, EDTA 10 mM, ARNasa A 100 µl/ml) y se lisa mediante la adición de 250 µl de tampón alcalino (NaOH 200 mM, SDS al 1%). Después de 5 min de homogeneización, se añaden 350 µl de tampón (acetato de potasio 3M, pH 5,5). Se transfiere a continuación el sobrenadante que contiene ADN plasmídico a una columna QIAprep. Una centrifugación elimina el eluido del tubo colector.

Se lava la columna con un tampón de etanol, se seca y se eluye entonces el ADN con 50 µl de agua estéril.

15 Para verificar la inserción del fragmento de interés, se digiere entonces el plásmido con la enzima de restricción EcoRI. Se efectúa la digestión en una mezcla de reacción de 30 µl que contiene: 2 µl de disolución plasmídica, 10 U de enzima EcoRI (Appligène) y tampón de reacción 1X. La digestión dura 2 horas a 37°C y el resultado se visualiza mediante electroforesis en gel de agarosa.

6.4- Secuenciación mediante el procedimiento de terminador BigDye

20 Se realiza la secuenciación con los productos de PCR previamente purificados en columnas Microcon 50 (Amicon) o con ADN plasmídico. Los fragmentos se secuencian directamente con los cebadores de PCR (fragmento R0 secuenciado con los cebadores 900S y 1280As) o bien después de clonación con el vector pCRII usando los cebadores universales (Sp6 y T7).

Se mantuvieron dos clones diferentes para cada una de las regiones amplificadas, con el fin de eliminar las eventuales ambigüedades en la lectura de las secuencias nucleotídicas.

25 Se efectúa la secuenciación con la ayuda del reactivo terminador BigDye (PE, Applied Biosystems). El principio de secuenciación consiste en una electroforesis vertical en gel de poliacrilamida del ADN marcado por cuatro fluorocromos diferentes. Se depositan las matrices de ADN sobre el gel y se separan según su tamaño, antes de someter el gel a un haz láser continuo. El láser excita los fluorocromos, que emiten cada uno a una longitud de onda diferente, detectada por un espectrógrafo. Un software informático, acoplado con el secuenciador, permite el análisis automático y la conversión de los datos en secuencias nucleotídicas.

30 La mezcla de reacción de 10 µl comprende: 4 µl de la solución de marcaje (trifosfatos de desoxinucleótidos (dATP, dCTP, dGTP, dUTP), ADN polimerasa AmpliTaq, MgCl₂, tampón Tris-HCl pH 9), 20 pmoles de cebador (codificante o anticodificante) y 500 ng de plásmido purificado sobre columnas de tipo Centricon. Se efectúan las reacciones de las secuencias (codificantes y anticodificantes) en un termociclador de tipo Perkin 9600, con 25 ciclos (96°C durante 10 s; 50°C durante 5 s; 60°C durante 4 min). Se precipitan a continuación los productos con 40 µl de etanol al 70%, se depositan sobre gel y se analizan gracias a un secuenciador automático de tipo ABI PRISM 377.

35 Se presentan las secuencias brutas obtenidas en forma de electroforogramas. La validación y explotación de las secuencias se hace con la ayuda del software Séquence Navigator (PE, Applied Biosystems). Son objeto de una doble lectura por dos cebadores de secuenciación diferentes (codificantes y anticodificantes), con el fin de minimizar los errores.

40 Se introducen a continuación estas secuencias en un ordenador usando el software DNA Strider 1.3, que permite un análisis rápido de las secuencias.

7- **Análisis informático de las secuencias nucleotídicas y proteicas**

Se comparan las secuencias leídas y corregidas y se someten a diferentes algoritmos filogenéticos.

45 Se compararon las secuencias obtenidas (22 secuencias) con 21 secuencias genómicas completas de HDV disponibles en GenBank (Tabla V).

Tabla V: Números de acceso a los diferentes aislamientos

Números de acceso (GenBank)	Nombre del aislamiento	Origen geográfico
1 X04451	Italia 1 (A20)	Italia
2 M84917	Líbano 1	Líbano
3 X85253	PacienteA	Cagliari (Italia)
4 X60193	7/18/83 (paciente S)	Japón
5 M92448	Taiwán	Taiwán
6 L22061	Colombia	Colombia
7 X77627	Suero humano chino	China central
8 L22064	Perú-2	Perú
9 L22063	Perú-1	Perú
10 L22066	US-2	Estados Unidos
11 M58629	Nauru	Isla de Nauru
12 U81988	Somalía	Somalía
13 U81989	Etiopía 1	Etiopía
14 AF098261	Canadá	Canadá (Québec)
15 U19598	Taiwán 3	Taiwán
16 AF018077	TW2b	Taiwán
17 L22062	Japón 3	Japón
18 AF309420	Miyako	Isla de Miyako (Okinawa)
19 D01075	US-1	Estados Unidos
20 M21012 W 15	W15	Transmisión experimental (marmota)
21 AJ307077	W5	Transmisión experimental (marmota)
22 AJ309868 a AJ309881	Aislamientos de Yakutia	Yakutia (Rusia)

5 La primera etapa consiste globalmente en alinear las secuencias de interés con las secuencias de HDV de referencia descritas y catalogadas en el banco de datos (Genbank), usando el programa CLUSTAL W1.8 (Thompson *et al.*, N.A.R., 1994, 22, 4673-4680). Han sido a veces necesarias correcciones manuales menores con la ayuda del software SeqPup con el fin de optimizar el alineamiento.

Se siguieron dos enfoques: el uso del alineamiento de las proteínas para el gen HD y el estudio de la estabilidad de las posiciones alineadas con la ayuda de un software de alineamiento adaptado.

10 A partir de este alineamiento de secuencias nucleotídicas, se construyen árboles filogenéticos mediante diferentes algoritmos. Los análisis están basados en matrices de distancia (enfoque fenético), cálculos de máxima parsimonia (MP; enfoque cladístico) y máximo parecido (ML, enfoque estadístico).

- Enfoque fenético (distancia genética)

15 El principio de este procedimiento es encontrar los pares de secuencias vecinas minimizando la longitud total de las ramas del árbol. Este enfoque permite reconstruir una filogenia basándose en el cálculo del parecido global entre las secuencias comparadas de dos en dos, lo que se expresa por una distancia. Es un procedimiento que permite convertir los datos de secuencias en valores numéricos de distancias, dispuestas en matriz. La topología del árbol está construida de manera que se agrupen las secuencias que poseen más caracteres en común, usando uno de los procedimientos de agrupamiento como el procedimiento del vecino más cercano o "neighbor-joining" (Saitou *et al.*, 1987).

- Enfoque cladístico (máxima parsimonia)

El principio de este procedimiento consiste en establecer relaciones de parentesco entre las secuencias mediante la búsqueda de bases nucleotídicas compartidas, minimizando los eventos genéticos. El algoritmo de máxima parsimonia construye un árbol filogenético de tal manera que implica el mínimo de mutaciones. El árbol mantenido es aquel que requiere menos cambios. Este procedimiento es sensible a las diferencias en las tasas de mutaciones a lo largo de las ramas. Los "clados" o "grupos monofiléticos" están constituidos por grupos de secuencias que comparten un ancestro común, con exclusión de cualquier otra secuencia.

- Enfoque estadístico

El procedimiento de máximo parecido se considera un enfoque estadístico. El programa calcula la probabilidad de que una secuencia evolucione hacia otra con el tiempo. En otros términos, consiste en considerar los cambios al nivel de cada sitio o carácter como eventos probabilísticos independientes. Este algoritmo de parecido se acumula para todos los sitios, y la suma se maximiza para estimar la longitud de la rama del árbol. Este procedimiento requiere un tiempo de cálculo grande para buscar el árbol filogenético, el más verosímil, correspondiente a las secuencias observadas debido a que tiene en cuenta la probabilidad de cambio de cada carácter.

Todos los análisis filogenéticos se realizaron usando los softwares informáticos Phylip 3.75 (PHYLogenetic Inference Package) (Felsenstein *et al.*, 1989) y Paup * versión 4.0 beta 6 (Phylogenetic Analysis Using Parsimony) (Swofford *et al.*, 1998).

Se calculó el análisis de la distancia mediante el procedimiento de Kimura con dos parámetros, que considera las tasas de transición (mutaciones T <-> C y G <-> A) en cada sitio, y de transversión (mutaciones "A o G! <-> "T o C") en cada sitio como diferentes.

La fiabilidad y robustez de estos agrupamientos de secuencias (o topologías) se evalúan estadísticamente mediante el enfoque de remuestra (o aleatorización muestral) sobre 10^3 y 10^4 remuestras.

Se presentan los resultados obtenidos en forma de un árbol filogenético visualizado por el programa Treeview (versión 1.6.5), que propone diferentes presentaciones del árbol (cladograma, radial y filograma). Permite visualizar igualmente los valores de aleatorización muestral en cada nodo y determinar un taxón como grupo externo (secuencias del genotipo III).

La traducción a aminoácidos del gen delta se efectúa con la ayuda del programa DNAStrider versión 1.3. El alineamiento de las secuencias proteicas se realiza como se describe anteriormente.

8- Análisis genotípico del HDV mediante polimorfismo de restricción (RFLP)

Se efectúa el genotipado de HDV mediante PCR-RFLP de la región R0 según las etapas siguientes:

- Etapas 1: Se digieren los productos de PCR por las dos enzimas de restricción SmaI y XhoI (New England Biolabs): se digieren 10 µl de producto amplificado separadamente en dos tubos con 5 U de enzimas SmaI o XhoI, respectivamente a 30°C y 37°C durante 3 horas en un volumen final de 50 µl en presencia del tampón apropiado y agua estéril. Se visualizan los productos de digestión bajo rayos ultravioleta como se describe anteriormente y se determinan los tamaños de los fragmentos mediante comparación con un marcador de tamaño (patrón de ADN de 50 pb, o los marcadores V y VI, Life Technologies GibcoBRL).

- Etapas 2: Se digieren las muestras que presentan un perfil distinto del perfil de genotipo I con otra enzima SacII (New England Biolabs) durante 3 horas a 37°C y se visualizan los productos de digestión como en la etapa 1.

- Etapas 3: Se determina el genotipo del virus a partir del análisis de la combinación de perfiles de restricción SmaI, XhoI y SacI.

9- Algoritmo de genotipado de HDV por PCR-RFLP

El algoritmo de genotipado de HDV por PCR-RFLP comprende al menos dos etapas:

- la primera consiste en el corte por dos enzimas de restricción, SmaI y XhoI, del fragmento R0 amplificado por PCR-TI a partir de ARN extraídos de sueros de pacientes;

- la segunda, para los pacientes de "perfil no I", consiste en un corte de R0 por la enzima SacII;

- la secuenciación de la región R0 o de la región que codifica p24 (o si es necesario de la totalidad del genoma), seguida de análisis filogenéticos, solo se realizará en segunda instancia si se obtienen perfiles de restricción inhabituales.

EJEMPLO 2: DESCUBRIMIENTO DE NUEVOS GENOTIPOS DE HDV

1- Análisis filogenético de la región R0

Se analizaron 22 muestras de pacientes infectados por HBV y HDV. Se amplificó la región R0 por PCR y se secuenció entonces el fragmento obtenido usando los cebadores 900S y 1280As.

Se efectuó el estudio filogenético a partir de un alineamiento de secuencias de 336 bases de R0 (las regiones ambiguas se eliminan) incluyendo, además de las 22 secuencias estudiadas, 15 secuencias de referencia y 6 secuencias R0 de Yakutia (Pt13, 26 (SEQ ID NO: 66), 29, 62 (SEQ ID NO:67), 63 y 704). El nombre dado a las secuencias corresponde a dFr (por "delta Francia") seguido del número de suero del paciente.

a) Análisis de la distancia genética

El árbol filogenético, obtenido mediante reconstrucción a partir de las distancias genéticas de la región R0, se presenta en la figura 1.

La topología del árbol individualiza los genotipos I y III, representados respectivamente por 7 y 3 secuencias nucleotídicas de referencia. Las demás secuencias de referencia se representan por las secuencias de tipo II (Japón, Taiwán-3 y las secuencias de Yakutia) y un grupo de dos secuencias (TW2b, Miyako) descritas respectivamente cada una como prototipos de los "subtipos IIB y IIC".

Este árbol muestra que las secuencias víricas procedentes de las 22 muestras analizadas corresponden a dos situaciones:

- 11 secuencias están asignadas a las secuencias de genotipo I, con la excepción de la secuencia dFr46, que parece estar emparentada con la secuencia US-1 descrita por Makino (Makino *et al.*, 1987); todas estas secuencias se distribuyen de forma heterogénea dentro del genotipo I;
- las 11 secuencias restantes están muy alejadas del genotipo I y del genotipo III. Además, ninguna de estas secuencias se agrupa directamente con las secuencias de tipo II (Japón, Taiwán-3, Yakutia 13, 26, 29, 62, 63, 704) o con el grupo de secuencias (TW2b, Miyako); estas secuencias de referencia forman por sí mismas dos grupos distintos.

La topología del árbol, obtenida mediante reconstrucción a partir de las distancias genéticas de la región R0, muestra que las moléculas nucleicas aisladas a partir de diversos HDV variantes se distribuyen en cuatro subgrupos (figura 1):

- la molécula dFr644 que parece aislada; posee no obstante, con un grupo de tres moléculas (dFr45, dFr2066 y dFr1843), un nodo mantenido por un valor de remuestra de solo un 66,7%;
- en contraposición, la rama que une las moléculas dFr45, dFr2066 y dFr1843 es robusta, ya que está mantenida por un valor de aleatorización muestral (BV) de un 99,9%.
- un conjunto de 5 moléculas: dFr47, dFr910, dFr69, dFr73 y dFr1953 está mantenido por un BV de un 100% y
- un par de moléculas dFr48 y dFr2020, que está igualmente mantenido por un BV de un 100%.

b) Análisis de máxima parsimonia

El árbol filogenético obtenido mediante reconstrucción a partir de la máxima parsimonia de la región R0, se presenta en la figura 2.

El análisis de la máxima parsimonia mantiene la misma topología que el análisis de la distancia genética. La reconstrucción evidencia la existencia en las 11 secuencias variantes de los mismos tres grupos monofiléticos: por ejemplo, con este enfoque el grupo de 5 moléculas dFr47, dFr910, dFr69, dFr73 y dFr1953 está igualmente mantenido por un BV de un 97% (Figura 2).

Las 11 moléculas variantes, las moléculas de genotipo II y el conjunto [TW2b, Miyako] parecen derivar de una rama común que podría individualizar, por comparación con las secuencias de genotipo I y genotipo III, el conjunto de secuencias de genotipo II. Sin embargo, los valores de aleatorización muestral que mantienen esta ramificación son relativamente modestos: un 88,5% en NJ y un 64,5% en MP (remuestreo efectuado sobre 10^4 muestras) en comparación con las del genotipo I (BV = 99,8%) y el genotipo III (BV= 100 %). Además, la distancia media entre los diferentes subgrupos definidos en los 11 HDV variantes, o entre estas variantes y las secuencias de genotipo II, parece más elevada que entre el conjunto de aislamientos de genotipo I o en las tres moléculas que definen el genotipo III.

El conjunto de estos resultados destaca la caracterización de nuevos genotipos de HDV.

2 – Análisis filogenético de la totalidad de los genomas

a) Reconstrucción del genoma completo a partir de fragmentos amplificados

Con el fin de estudiar el genoma completo de estas variantes, y con el objetivo de precisar su asignación, se amplificaron varias regiones del genoma de HDV (Tabla II) a partir de 6 muestras que incluyen al menos un miembro de cada uno de los 4 subgrupos, y se seleccionaron 3 representantes del grupo mayoritario: dFr45, dFr47, dFr48, dFr73, dFr644 y dFr910.

De manera más precisa, se amplificaron los fragmentos siguientes por PCR (Tabla IV):

- fragmentos de 850 pb (R'1) y 1050 pb (R'2) superpuestos en sus extremos para dFr45, dFr47, dFr48 y dFr73,
- dos fragmentos superpuestos de 960 pb y de 1250 pb para dFr644, y
- dos fragmentos de 1400 pb y 650 pb para dFr910.

5 Se clonaron todas estas regiones genómicas amplificadas en un vector pCRII™ (Tabla VI). Se secuenciaron dos clones correspondientes a cada uno de los fragmentos amplificados. Se efectuó la reconstrucción de secuencias completas de consenso de ADNc de HDV después del alineamiento de las regiones superpuestas y el alineamiento con las secuencias de referencia.

Tabla VI: Clones de pCRII que contienen los diferentes insertos

	R0	R'1	R'2	G910	R1
dFr45	-	dFr45R'1 clon 2 dFr45R'1 clon 4	dFr45R2 clon 8 dFr45R2 clon 10	-	-
dFr47	-	dFr47R'1 clon 13 dFr47R'1 clon 16	dFr47R2 clon 19 dFr47R2 clon 22	-	-
dFr48	-	dFr48R'1 clon 23 dFr48R'1 clon 28	dFr48R2 clon 19 dFr48R2 clon 22	-	-
dFr73	-	dFr73R'1 clon 36 dFr73R'1 clon 39	dFr73R2 clon 29 dFr48R2 clon 33	-	-
dFr644	-	-	-	-	644R1 clon 4 644R1 clon 8
dFr910	910R0-clon 4 910R0-clon	-	-	R910 clon 29 R910 clon 31	910R1 clon 4 910R1 clon 5

10

b) Análisis de 6 nuevas secuencias genómicas completas de HDV de origen africano

b1) Características clínicas de los 6 pacientes

5 pacientes son originarios de África Occidental, y un paciente ha estado en Camerún. En el momento de la toma de muestra, estos pacientes residían en la región parisina desde hacía al menos dos años. Todos estos pacientes estaban aquejados de hepatitis grave y los datos clínicos se resumen en la figura 3.

15

b2) Organización genómica de las nuevas secuencias de HDV

El análisis comparativo de las regiones R0 de 22 pacientes infectados por HDV y HBV con aquellas disponibles en las bases de datos ha evidenciado la gran diversidad genética del genoma vírico de HDV.

El tamaño de los genomas completos es diferente para las 6 secuencias de los 6 aislamientos de HDV de origen africano, lo que confirma la variabilidad del HDV:

20

- el genoma vírico de los aislamientos dFr910, dFr47 y dFr73, que comprende 1697 nucleótidos, es el más largo descrito nunca para HDV;
- el genoma del aislamiento dFr45 aparece entre los más pequeños (1672 nt), y
- las secuencias genómicas de los virus dFr644 y dFr48 son respectivamente de 1680 nt y de 1687 nt.

25 El análisis después del alineamiento de las diferentes secuencias estudiadas revela un grado elevado de conservación en las regiones del genoma de HDV correspondientes a las ribozimas responsables de las escisiones de los ARN genómicos y antigenómicos. De igual modo, el marco de lectura que codifica el antígeno delta se encuentra en la hebra antigenómica. Un codón de triptófano (UGG) es el único caracterizado por dos secuencias (dFr47, dFr910) y una ambigüedad (G/A) encontrada para las otras cuatro secuencias indica que la proteína delta pequeña y grande son muy

probablemente sintetizadas. Las regiones variables comprenden la porción no codificante así como los extremos 5' y 3' del gen LHD. Notablemente, existe una inserción de 7 nucleótidos en la secuencia dFr48. Esta inserción está presente en un bucle correspondiente a uno de los extremos del genoma en su forma pseudocatenaria (en la posición 797 de la secuencia de referencia Italia (Wang *et al.*, 1987).

5 c) Comparación de las 6 secuencias de HDV de origen africano con las secuencias representativas de los diferentes genotipos

La comparación de las 6 nuevas moléculas con las moléculas conocidas, representativas de los tres genotipos conocidos, indica una similitud de nucleótidos situada entre 71,7% (dFr45 frente a Libano) y 80,0% (dFr73 frente a Yakutia p26) con respecto a las moléculas de genotipos I, II y las moléculas TW2b y Miyako. Efectivamente, para cada uno de los 6 aislamientos, la media de similitud en nucleótidos es del orden de 73,3% a 74,6% con las moléculas del genotipo I, de 74,5% a 78,8% con las del genotipo II y del orden de 74,6% a 77,8% con las moléculas TW2b/Miyako. En contraposición, la similitud nucleotídica con el aislamiento de Perú (genotipo III) solo es de 63,9% a 66,0%, confirmando el alejamiento particular de esta molécula (Tabla VII). Además, cuando se comparan entre ellas, las 6 moléculas correspondientes a estos genomas completos y que definen las 6 variantes dFr45, dFr47, dFr48, dFr73, dFr644 y dFr910, solo el grupo de moléculas dFr73, dFr910 y dFr47 presenta una similitud de secuencia del orden de un 90%. Las moléculas dFr45, dFr48 y dFr644 están tan alejadas entre sí como de los genotipos I, II, de las secuencias TW2b/Miyako y del grupo de moléculas dFr73, dFr910 y dFr47 (del orden de 73,2% a 78%) (Tabla VIII).

Tabla VII: Porcentaje de similitud de las secuencias completas de HDV africanos con los diferentes genotipos conocidos (cálculo de la media)

Aislamientos de HDV*	Tipo I	Tipo II	TW2b/Miyako	Tipo III
dFr45	73,3 71,7 - 74,6	74,5 73,2 - 75,5	74,6	66
dFr47	74,2 73,0 - 75,0	78,6 78,2 - 79,9	77,4	65,5
dFr48	73,3 72,0 - 74,0	77,1 76,6 - 77,7	75,5 74,4 - 76,6	65,4
dFr73	74,1 73,0 - 75,0	78,8 77,7 - 80,0	77,8 77,5 - 78,0	65,9
dFr644	73,6 72,2 - 74,6	76,8 76,2 - 77,2	77,0 76,9 - 77,2	63,9
dFr910	74,6 73,0 - 75,8	77,9 77,0 - 78,6	77,2 77,0 - 77,5	64,6

*Los aislamientos de HDV de referencia corresponden a los genomas completos estudiados en el ejemplo 2.1

Tabla VIII: Porcentaje de similitud de las nuevas moléculas de HDV entre ellas

	<i>dFr47</i>	<i>dFr48</i>	<i>dFr73</i>	<i>dFr644</i>	<i>dFr910</i>
<i>dFr45</i>	74,8	73,2	75	78	74,7
<i>dFr47</i>		77,1	90	76,3	89
<i>dFr48</i>			77,7	75,5	76,1
<i>dFr73</i>				76,3	89
<i>dFr644</i>					76,1

d) Análisis filogenético de 6 moléculas de HDV de origen africano y de moléculas representativas de los diferentes genotipos

Se llevó a cabo el análisis filogenético en las 6 secuencias completas de origen africano, 16 secuencias de referencia y 2 secuencias de Yakutia (Pt26 y Pt62). La Figura 4 ilustra los resultados obtenidos mediante análisis de distancia. El árbol filogenético reconstruido por "Neighbor-Joining" (NJ) muestra que ninguna de las 6 secuencias estudiadas (dFr45, dFr47, dFr48, dFr73, dFr644 y dFr910) se agrupa con las secuencias de referencia del genotipo I o del genotipo III. La asignación de estas secuencias africanas a las secuencias del genotipo II (a las secuencias TW2b y Miyako descritas respectivamente como los subtipos IIB y IIC) no está apoyada por valores de aleatorización muestral elevados (<70 %) (Wu *et al.*, 1998). Además, las secuencias TW2b y Miyako parecen formar un grupo distinto y monofilético con un BV de 100%. Estas dos secuencias parecen constituir por sí mismas un "clado" que representa un genotipo diferente del de tipo II.

En los análisis de distancia, las 6 secuencias africanas se subdividen en 3 subgrupos distintos (apoyados por BV superiores a 90,3% para 10^4 remuestras). Las secuencias dFr47, dFr73 y dFr910 constituyen un grupo cuya rama se basa en un valor de aleatorización muestral de 100%. Para apoyar estos resultados, se realizó el estudio de máxima parsimonia con el mismo conjunto de secuencias (Figura 5). Al enraizar el árbol de forma artificial gracias a la secuencia "Perú-1", se individualiza como en todos los análisis llevados a cabo anteriormente el conjunto de secuencias del genotipo I (BV = 100 %). La topología de las demás secuencias mantiene la distribución de aislamientos africanos y asiáticos en varios grupos; esto muestra el interés del uso de la región R0. El genotipo II agrupa las secuencias de Yakutia, Taiwán-3 y Japón con un BV de 99,9% sobre 10^4 remuestras. Igualmente, se confirma la individualización de TW2b y Miyako (BV = 100%). Por último, las secuencias africanas indican la existencia de al menos 3 subgrupos. La monofilia de las secuencias dFr47, dFr73 y dFr910 (BV = 100 %) mantiene la asignación de estas secuencias a un subgrupo. En contraposición, la secuencia dFr48, que posee con los aislamientos del grupo precedente (dFr910, dFr47, dFr73) una similitud nucleotídica respectiva de 76,1, 77,1 y 77,7%, se agrupa con estas secuencias solo en un 55,4% de las remuestras, sugiriendo que es posible su individualización. Aunque parecen distantes entre sí, se observa el agrupamiento de dFr45 y dFr644, con un BV elevado (NJ = 96,5 / MP = 88,6) en el contexto estudiado.

Por lo consiguiente, los análisis filogenéticos de las regiones R0 y de las secuencias completas de las secuencias africanas indican a la vez que los grupos difieren entre sí y podrían constituir 3 (hasta 4) genotipos distintos; estos resultados evidencian así la existencia de al menos 7 genotipos de HDV.

3- Análisis de la secuencia aminoacídica (aa) del antígeno delta (Ag-HD).

El Ag-HD está representado por las dos formas p24 (sHD) y p27 (LHD) de la proteína delta. La secuencia proteica de 1 a 194-195 aminoácidos corresponde a la proteína delta pequeña (sHD) o forma p24. La proteína delta grande (LHD) o forma p27 posee el mismo extremo aminoterminal y una extensión de 19 a 20 aminoácidos en su extremo carboxiterminal.

El alineamiento de la secuencia del antígeno HD de las 6 secuencias africanas con las secuencias de las proteínas HD conocidas se presenta en la figura 6.

El análisis de las secuencias muestra que los 6 aislamientos de origen africano tienen una identidad aminoacídica del orden de 69 a 77% con las secuencias del genotipo I, de 71 a 79% con los aislamientos del genotipo II, de 72 a 78% con las secuencias TW2b/Miyako y de 63% con el aislamiento de Perú (genotipo III).

El tamaño de las proteínas correspondientes a los nuevos aislamientos varía entre 213 y 214 aminoácidos. Todas estas proteínas tienen el mismo perfil de hidrofobicidad. La forma p24 posee 2 regiones hidrófobas pequeñas, una situada alrededor de los aminoácidos 50-60 (entre el sitio de polimerización y el SLN) y la otra entre las posiciones 160 y 172 (en consideración de un motivo extremadamente conservado). Otros dos dominios están bien conservados entre los diferentes genotipos: se trata del dominio de fijación a ARN y del dominio de localización nuclear. A semejanza de lo descrito en la bibliografía, el extremo carboxiterminal de la proteína delta (entre los aminoácidos 195 y 215) constituye una región hipervariable. Solo dos aminoácidos de los 19-20 están conservados. Se trata de la cisteína (C) correspondiente al sitio de farnesilación de la forma grande de la proteína HD y la glicina (G) carboxiterminal. Además, se encuentran las secuencias señal específicas de aislamientos del mismo genotipo, por ejemplo, los 19 aminoácidos específicos de la proteína grande del genotipo I o los 20 aminoácidos del genotipo III.

En contraposición, para las secuencias proteicas de los aislamientos de origen africano, los aislamientos de genotipo II y TW2b/Miyako, el extremo carboxiterminal parece subdividido en dos dominios. El dominio variable está representado por los aminoácidos 197 a 205 y el dominio conservado va de los aminoácidos 206 a 215 (RLPLLECTPQ) (Figura 5).

4- Definiciones de los 7 clados de HDV

El análisis de las secuencias completas de los 6 aislamientos de África permite definir 7 clados de HDV correspondientes a los genotipos siguientes (Tabla IX):

Tabla IX: Correspondencia clado/genotipo

Clado	Genotipo	Aislamientos
1	I	Italia, W5, W15, Us 1, Us2, Líbano, Etiopía, Somalia, isla de Nauru, China, Cagliari, Canadá...
2	IIA	Japón, Taiwán ³ , Yakutia ²⁶ , Yakutia ⁶²
3	III	Perú ¹
4	IIB, IIC	TW2b, Miyako
5	?	dFr910, dFr73, dFr47
6	?	dFr48
7	?	dFr45, dFr644

EJEMPLO 3: PROCEDIMIENTO DE GENOTIPADO DE HDV-1 a HDV-7 por PCR-RFLP

Se efectúa el genotipado según el protocolo descrito en el ejemplo 1.8.

5 1- Falta de sensibilidad de PCR 6A/6S

Inicialmente, tres pacientes positivos de Ag-HBs plantearon un problema de diagnóstico de infección delta. Efectivamente, en estos pacientes se observa una hepatitis grave asociada a la presencia de IgM anti-HDV, pero una ausencia de replicación de HDV por PCR-TI con la ayuda de los cebadores 6A-6S descritos en Deny *et al.* (1991, 1993, 1994, anteriormente citados) para el diagnóstico rutinario de infección por HDV. La PCR 6A/6S amplifica un fragmento de ADNc de 234 pb correspondiente a la parte del extremo carboxiterminal del gen LHD (posiciones 904 a 1141 en el genoma vírico).

Se volvieron a amplificar los ARN extraídos del suero de estos mismos pacientes con la ayuda del par de cebadores 900S y 1280AS que definen la región R0.

Los resultados obtenidos a partir de muestras de estos tres pacientes han evidenciado la presencia reproducible de una banda de 400 pb (R0) con los cebadores 900S y 1280AS, mientras que la PCR 6A-6S seguía siendo negativa.

Se confirmaron estos resultados con una serie de muestras de suero de pacientes que se analizaron en paralelo con los pares de cebadores 6A-6S y 900S-1280AS. De 286 muestras, 14 son positivas únicamente con PCR R0.

Estos resultados demuestran una mayor especificidad y una mejor sensibilidad de los cebadores 900S y 1280AS, en comparación con los cebadores 6S y 6A, para la detección del ARN de HDV en el suero de pacientes infectados.

20 2- Perfiles de restricción esperados para HDV-1 a HDV-7

Los procedimientos usados clásicamente de PCR-RFLP (Wu *et al.*, 1995a; Wu *et al.*, 1995b; Casey *et al.*, 1996) permiten la distinción de tres genotipos delta diferentes. El uso de la enzima de restricción SmaI no diferencia todos los genotipos I, IIA, IIB reconocidos hasta ahora, y se ha usado la enzima XhoI para diferenciar el "subtipo IIA" del "subtipo IIB" (Wu *et al.*, 1995b).

La asociación de las dos enzimas SmaI y XhoI en una primera etapa revela 7 perfiles distintos (de P1 a P7) (Tabla X). Estos 7 perfiles no se superponen exactamente con los 7 clados (HDV-1 a HDV-7). En consecuencia, se escinden las muestras de perfil "no P1" en una segunda etapa con la enzima SacII, desembocando así en la obtención, de forma combinada, de 10 perfiles delta distintos (de D1 a D10) (Tabla XI) que pueden vincularse específicamente con los diferentes clados descritos gracias al análisis de filogenia.

Tabla X: Perfiles de corte por restricción de la región R0 por las enzimas SmaI y XhoI

ETAPA 1 Genotipos descritos	Fragmentos de SmaI Tamaño (pb)	Perfil de SmaI	Fragmentos de Tamaño (pb)	XhoI	Perfil de XhoI	Perfil combinado de SmaI-XhoI	
I	220, 179	S1	383, 16	X1	S1 X1		P1
IIA	397	S2	303, 78, 16	X2	S2 X2		P2
IIIB	397	S2	319, 79	X3	S2 X3		P3
IIC (Miyako)	397	S2	157, 162, 79	X4	S2 X4		P4
III	298, 107	S3	405	X5	S3 X5		P5
II Yakutia	178, 117, 110	S4	303, 78, 16	X2	S4 X2		P6
dFr45	217, 179	S1	303, 78, 16	X2	S1 X2		P7
dFr644	217, 179	S1	303, 78, 16	X2	S1 X2		P7
dFr47, 73, 910	179, 111, 107	S4	303, 78, 16	X2	S4 X2		P6
dFr48	397	S2	303, 78, 16	X2	S2 X2		P2

Tabla XI: Perfiles de restricción esperados después del corte de la región R0 por las enzimas SmaI, XhoI y SacII

ETAPA 2 Genotipos descritos	Fragmentos de SacII Tamaño (pb)	Perfil de SacII	Perfil combinado de SmaI-XhoI/SacII	
I	362, 38	Sc1	S1 X1 Sc1	D1
IIA	266, 92, 38	Sc2	S2 X2 Sc2	D2
IIIB	268, 130	Sc3	S2 X3 Sc3	D3
IIC (Miyako)	268, 130	Sc3	S2 X4 Sc3	D4
III	405	Sc4	S3 X5 Sc4	D5
II Yakutia	266, 92, 38	Sc2	S4 X2 Sc2	D6
dFr45	268, 130	Sc3	S1 X2 Sc3	D7
dFr644	397	Sc4	S1 X2 Sc4	D8
dFr47, 73, 910	268, 130	Sc3	S4 X2 Sc3	D9
dFr48	268, 130	Sc3	S2 X2 Sc3	D10

3- Genotipado de muestras de pacientes por PCR-RFLP

A partir del análisis por PCR-RFLP de muestras (más de 50):

- 5 - no se encontró ningún genotipo II ni III;
- un 89,7% de los pacientes presentaban un perfil D1 (genotipo I) y un 10,3% un perfil “no I”;
- se detectaron 2 nuevos perfiles con XhoI (X6 y X7) que conllevan tres nuevas combinaciones (D11, D12 y D13) suplementarias (Tablas XII y XIII).

Tabla XII: Nuevos perfiles de restricción de XhoI obtenidos en 5 pacientes originarios de África Occidental

ETAPA 1 Pacientes	Fragmentos de SmaI Tamaño (pb)	Perfil de SmaI	Fragmentos de XhoI Tamaño (pb)	Perfil de XhoI	Perfil combinado de SmaI-XhoI	
dFr1843	218, 179	S1	303, 78, 16	X2	S1 X2	P7
dFr1953	218, 179	S1	303, 78, 16	X2	S1 X2	P7
dFr2020	392	S1	303, 73, 16	X2	S2 X2	P2
dFr2088	220, 179	S1	242, 171, 16	X6	S1 X6	P8
dFr2066	217, 179	S1	237, 66, 16	X7	S1 X7	P9

10

Tabla XIII: Nuevos perfiles de restricción de XhoI, SmaI, SacII obtenidos en 5 pacientes originarios de África Occidental

ETAPA 2 Pacientes	Fragmentos de SacII Tamaño (pb)	Perfil de SacII	Perfil combinado de SmaI-XhoI/SacII	
dFr1843	267, 130	Sc3	S1 X2 Sc3	D7
dFr1953	267, 92, 38	Sc2	S1 X2 Sc2	D11
dFr2020	262, 130	Sc3	S2 X2 Sc3	D10
dFr2088	396	Sc4	S1 X6 Sc4	D12
dFr2066	396	Sc4	S1 X7 Sc4	D13

Se presenta la correspondencia entre los perfiles combinados y los genotipos identificados por el análisis filogenético en la Tabla XIV.

Tabla XIV: Resumen de los diferentes resultados basados en los análisis filogenéticos y los diferentes perfiles correspondientes

Clados	Genotipos	Aislamientos	Perfiles combinados (SmaI-XhoI / SacII)
HDV-1	I	Italia dFr2088	D1A D1B
HDV-2	IIA	Japón Aislamientos de Yakutia	D2A D2B
HDV-3	III	Perú1	D3
HDV-4	IIB IIC	TW2b Miyako	D4A D4B
HDV-5	V	dFr47, dFr73 y dFr910 dFr1953	D5A D5B
HDV-6	VI	dFr48, dFr2020	D6
HDV-7	VII	dFr45, dFr1843 dFr2066 dFr644	D7A D7B D7C

5

REFERENCIAS BIBLIOGRÁFICAS

- Casey J.L. *et al.*, Proc. Natl. Acad. Sci. USA, 1993a, 90, 9016-20.
- Casey J.L. *et al.*, J. Infect. Dis., 1996b, 174, 920-6.
- Chang F.L. *et al.*, Proc. Natl. Acad. Sci. USA, 1991, 88, 8490-8494.
- 10 • Chao Y.C. *et al.*, Hepatology, 1991b, 13, 345-52.
- Deny P. *et al.*, Res. Virol., 1994, 145, 287-95.
- Deny P. *et al.*, J. Med. Virol., 1993, 39, 214-8.
- Deny P. *et al.*, J. Gen. Virol., 1991, 72, 735-9.
- Felsenstein J. *et al.*, Cladistics, 1989, 5, 164-166.
- 15 • Gaeta G.B. *et al.*, Hepatology, 2000, 32, 824-7.
- Glenn J.S. *et al.*, Science, 1992, 256, 1331-3.
- Hwang S. *et al.*, Virology, 1993a, 193, 924-931.
- Imazeki F. *et al.*, J. Virol., 1990, 64, 5594-5599.
- Imazeki F. *et al.*, Nucl. Acid. Res., 1991, 19, 5439-5440.
- 20 • Lai M.M.C. *et al.*, S. Hadziyannis, J. Taylor y F. Bonino (ed.), "Hepatitis delta virus. Molecular biology, Pathogenesis, and Clinical aspects", 1993, 382, 21-27. Wiley-Liss, Nueva York.
- Lee C.M. *et al.*, Virology, 1992, 188, 265-273.
- Lee C.M. *et al.*, J. Med. Virol., 1996b, 49, 145-54.

- Makino S. *et al.*, Nature, 1987, 329, 343-6.
- Nakano T. *et al.*, J. Gen. Virol., 2001, 82, 2183-2189.
- Niro G.A. *et al.*, J. Hepatol., 1999, 30, 564-9.
- Niro G.A. *et al.*, Hepatology, 1997, 25, 728-34.
- 5 • Roingeard P. *et al.*, Clin. Infect. Dis., 1992, 14, 510-14.
- Saitou N. *et al.*, Mol. Biol. Evol., 1987, 4, 406-25.
- Sakugawa H. *et al.*, J. Med. Virol., 1999, 58, 366-72.
- Shakil A.O. *et al.*, Virology, 1997, 234, 160-7.
- Swofford D. *et al.*, PAUP*: "Phylogenetic Analysis Using Parsimony (and other methods), version 4.0d64".
- 10 • Thompson J.D. *et al.*, Nuc. Acid. Res., 1994, 22, 4673-80.
- Wang K.S. *et al.*, Nature, 1986, 323, 508-514.
- Wang K.S. *et al.*, Nature, 1987, 328, 456.
- Wu J.C. *et al.*, Hepatology, 1995a, 22, 1656-60.
- Wu J.C. *et al.*, J. Gen. Virol., 1998, 79, 1105-13.
- 15 • Wu J.C. *et al.*, Lancet, 1995b, 346, 939-41.

LISTADO DE SECUENCIAS

- <110> ASSISTANCE PUBLIQUE-HOPITAUX DE PARIS
DENY Paul
RADJEF Nadjia
20 HUC-ANAIS Patricia
- <120>
<130> S1020FR17
<140>
<141>
- 25 <160> 67
<170> PatentIn ver. 2.1
<210> 1
<211> 1672
<212> ADN
- 30 <213> virus de la hepatitis D
<220>
<221> fuente
<222> (1)..(1672)
<223> secuencia genómica completa de dFr45
- 35 <400> 1

```

atgagccaag tccgacgaag agtccggaga tgggagagga gagagacccc cgaggaggca 60
aagtcgagag ggcggagaca ctctgaggag aggactccca agaggacaag aaggactcaa 120
gatccgggaa cgaatcccca aaacgccggt aaagctcttg gaatgaggaa aacgggcggg 180
atggtagaaa agagcgagcc tcccgattac gagattcccc tgaacttatt ggatctggag 240
actccgacct taaggggtca aaggtaggac agggaggcaa tccaccagga gatttcggac 300
aactcacctc cagaggaccc cttcagcgaa cagaagagcc tgtcccggtc gaattgtccc 360
gtagcgatag gaggagatgc taggagtagg gagagaccgg tgcgagagag gcaagacaag 420
agagcagcgg ggctagcgag tggatgttcc gccccccgtt ggctccgagt gaggcttatc 480
ccggggaatt tggctawctt cccaactag ccggtcccgg atccccctcc aratgaacgg 540
ggggactccg gcctctcggg gaatcggcgg ctccatggtg aacccccgca gcccccact 600
ccacactcct tcccccgcg gggccccca tccaaagatg gaactccacc cctcagggtt 660
cgcyatccta cccttttctt acctttggcc ggcatggtcc cagcctcctc gctggcggcg 720
gctgggcaac attccgaggg gaaccgtccct cggcaatggc gaatgggacc caagacctcc 780
aagattccca gagagaatct ggaggtgact gggcactccc tttgccatcc gagtggacgt 840
tcgtcctcct tcggatgccc aggtcggacc acggggaggt ggagatgcca tgccgaccgg 900
aagaggaaag aaggattcgg acgcaaact gtgagtggaa gttcgtctt tattgggggg 960
tacactcgag gagtggaagg cggggagggg gggtcggact aggtccctat ggaaactgtc 1020
ggctcctcgc gatgtcgagt ccctctcccg ttctggagaa gggggactcc gggacacctc 1080
gcagttgagg aacgaagccg ccccgggcg ctccccctcg tggccctcgc ggagggttca 1140
catcccaac ccgcgggccc gctattcttc ttgctcttg ctgctcatcg gcggtcaact 1200
tcctgagttc ctcttcttc tccttgctga gggactttcc tcctgcgac agctgcttct 1260
tcttgttctc gagggccttc ctctgctcgg gagcccgtct ctctcgtcgc gtgaagctc 1320
ccttgttact cttttccct ggtccggagt cgacctccat ccgatctgtt cggggtctct 1380
tcgcccgggg agctccctcc ccagtcccga cttttccaat tattccttg atgtttccca 1440
gccagggatt gtcgtcctca agtttcttga tggcttctt tgtcttccgg agcctcctct 1500
cgagttcttc cgcgtcttcc ctgctttgw cccacttcga aagggctcct tcccttccct 1560
tccgggatct ctctgatcgc gcgtggctca tctcggcaga ggcgggcgat cctcagttct 1620
cttactcttt tcttgaaaga ggagactgct ggacgctccg ccgtgtccga gc 1672

```

<210> 2

<211> 642

<212> ADN

5 <213> virus de la hepatitis D

<220>

<221> CDS

<222> (1)..(642)

REIVINDICACIONES

1. Moléculas de ácido nucleico aisladas caracterizadas porque se seleccionan del grupo constituido por:
 - el genoma completo de la variante de HDV denominada dFr48, que presenta la secuencia SEQ ID NO: 21, y
 - el genoma de un HDV que presenta una divergencia genética $\leq 15\%$ con la secuencia SEQ ID NO: 21.
- 5 2. Moléculas de ácido nucleico aisladas, caracterizadas porque comprenden al menos uno de los fragmentos de las secuencias de una variante HDV según la reivindicación 1, seleccionados del grupo constituido por:
 - a) el fragmento R0, que presenta la secuencia SEQ ID NO:50,
 - b) el fragmento R1, que se extiende de las posiciones 307 a 1289 del genoma de HDV,
 - c) el fragmento R2, que se extiende de las posiciones 889 a 328 del genoma de HDV,
 - 10 d) el fragmento R3, que se extiende de las posiciones 1486 a 452 del genoma de HDV,
 - e) el fragmento R'1, que se extiende de las posiciones 305 a 1161 del genoma de HDV,
 - f) el fragmento R'2, que se extiende de las posiciones 984 a 331 del genoma de HDV,
 - g) el fragmento R644, que se extiende de las posiciones 889 a 446 del genoma de HDV,
 - h) el fragmento G910, que se extiende de las posiciones 1206 a 929 del genoma de HDV,
 - 15 i) el fragmento p910, que se extiende de las posiciones 553 a 1550 del genoma de HDV,
 - j) el ADNc que codifica la proteína sHD, de secuencia SEQ ID NO:24, y
 - k) el ADNc que codifica la proteína LHD, de secuencia SEQ ID NO:22.
3. Método de detección de una variante HDV tal como se define en la reivindicación 1, mediante hibridación y/o amplificación realizada a partir de una muestra biológica, estando caracterizado dicho método porque comprende:
 - 20 (1) una etapa de extracción del ácido nucleico a detectar, que pertenece al genoma del virus eventualmente presente en la muestra biológica,
 - (2) la realización de al menos una amplificación génica con la ayuda de un par de cebadores seleccionados del grupo constituido por los cebadores capaces de amplificar una de las regiones siguientes del ARN genómico de HDV: R0, R1, R2, R3, R644, G910, p910, R'1 y R'2 y
 - 25 (3) el análisis del producto amplificado por comparación con la molécula de secuencia SEQ ID NO:21
4. Método de detección según la reivindicación 3, caracterizado porque la etapa (3) de análisis puede ejecutarse mediante restricción, secuenciación o hibridación.
5. Método según la reivindicación 3 o la reivindicación 4, caracterizado porque los cebadores específicos para la amplificación de las regiones R0, R1, R2, R3, R644, G910, p910, R'1 y R'2, ejecutada en la etapa (2), se seleccionan del grupo constituido por:
 - 30 - los cebadores 900S (SEQ ID NO:33) y 1280AS (SEQ ID NO:34) para la amplificación de R0 (aproximadamente 400 pb),
 - los cebadores 320S (SEQ ID NO:39) y 1280AS (SEQ ID NO:34) para la amplificación del fragmento R1 (aproximadamente 960 pb),
 - 35 - los cebadores 900S (SEQ ID NO:33) y 320AS (SEQ ID NO:45) para la amplificación de R2 (aproximadamente 1100 pb), que contiene el gen sHD correspondiente a las posiciones 1598-950,
 - los cebadores 1480S (SEQ ID NO:46) y 440AS (SEQ ID NO:47) para la amplificación de R3 (aproximadamente 650 pb),
 - los cebadores 318S (SEQ ID NO:35) y 1150AS (SEQ ID NO:36) para la amplificación de R'1 (aproximadamente 850 pb),
 - 40 - los cebadores 960S (SEQ ID NO:37) y 345AS (SEQ ID NO:38) para la amplificación de R'2 (aproximadamente 1050 pb).

6. Método de detección y de genotipado de HDV a partir de una muestra biológica, estando caracterizado dicho método porque comprende:
- (a) una etapa de extracción del ácido nucleico que pertenece al genoma del virus HDV,
 - (b) una etapa de amplificación de la región R0 delimitada por las posiciones 889 a 1289 del genoma de HDV,
 - 5 (c) un primer tratamiento de las moléculas de ácidos nucleicos amplificadas con las enzimas de restricción SmaI y XhoI, produciendo un primer conjunto de fragmentos de restricción, y
 - (d) un segundo tratamiento de las moléculas de ácidos nucleicos por la enzima de restricción SacII, produciendo un segundo conjunto de fragmentos de restricción,
 - 10 (e) el análisis combinado de los dos conjuntos de fragmentos de restricción producidos por RFLP (polimorfismo de longitud de fragmento de restricción), de manera que se detecte la presencia del HDV dFr48 presente en dicha muestra biológica.
7. Método según la reivindicación 6, caracterizado porque la etapa (b) de amplificación de moléculas de ácido nucleico de dicha muestra por PCR-TI se ejecuta ventajosamente con los cebadores 900S (SEQ ID NO: 33) y 1280AS (SEQ ID NO:34).
- 15 8. Método según la reivindicación 6 o la reivindicación 7, caracterizado porque comprende además:
- (f) la amplificación de moléculas de ácido nucleico de dicha muestra por PCR-TI con los cebadores 900S (SEQ ID NO:33) y 320AS (SEQ ID NO: 45) de manera que se amplifique la región R2 y
 - (g) la secuenciación directa de la región R2 amplificada y la comparación con la molécula de ARN de secuencia SEQ ID NO: 21.
- 20 9. Vector recombinante, especialmente un plásmido, que comprende un inserto constituido por una molécula de ácido nucleico según cualquiera de las reivindicaciones 1 o 2.
10. Célula transformada por una molécula de ácido nucleico según una cualquiera de las reivindicaciones 1 o 2.
11. Productos de traducción codificados por la molécula de ARN de secuencia SEQ ID NO: 21, por sus secuencias complementarias codificantes o anticodificantes.
- 25 12. Proteínas caracterizadas porque están codificadas por una de las moléculas de ácido nucleico según cualquiera de las reivindicaciones 1 o 2.
13. Proteína según la reivindicación 12, caracterizada porque se selecciona del grupo constituido por:
- la proteína LHD de dFr48 que presenta la secuencia SEQ ID NO: 23, y
 - la proteína sHD de dFr48 que presenta la secuencia SEQ ID NO: 25.
- 30 14. Péptido caracterizado porque está constituido por un fragmento de una proteína según la reivindicación 12 o la reivindicación 13, seleccionado del grupo constituido por:
- el péptido A constituido por los 19 aminoácidos del extremo carboxiterminal de la secuencia SEQ ID NO: 23,
 - el péptido C de secuencia SEQ ID NO: 64.
- 35 15. Uso de una molécula de ácido nucleico según cualquiera de las reivindicaciones 1 o 3, o de una proteína según cualquiera de las reivindicaciones 11 a 14 para la preparación de un kit de detección y de genotipado de un HDV.

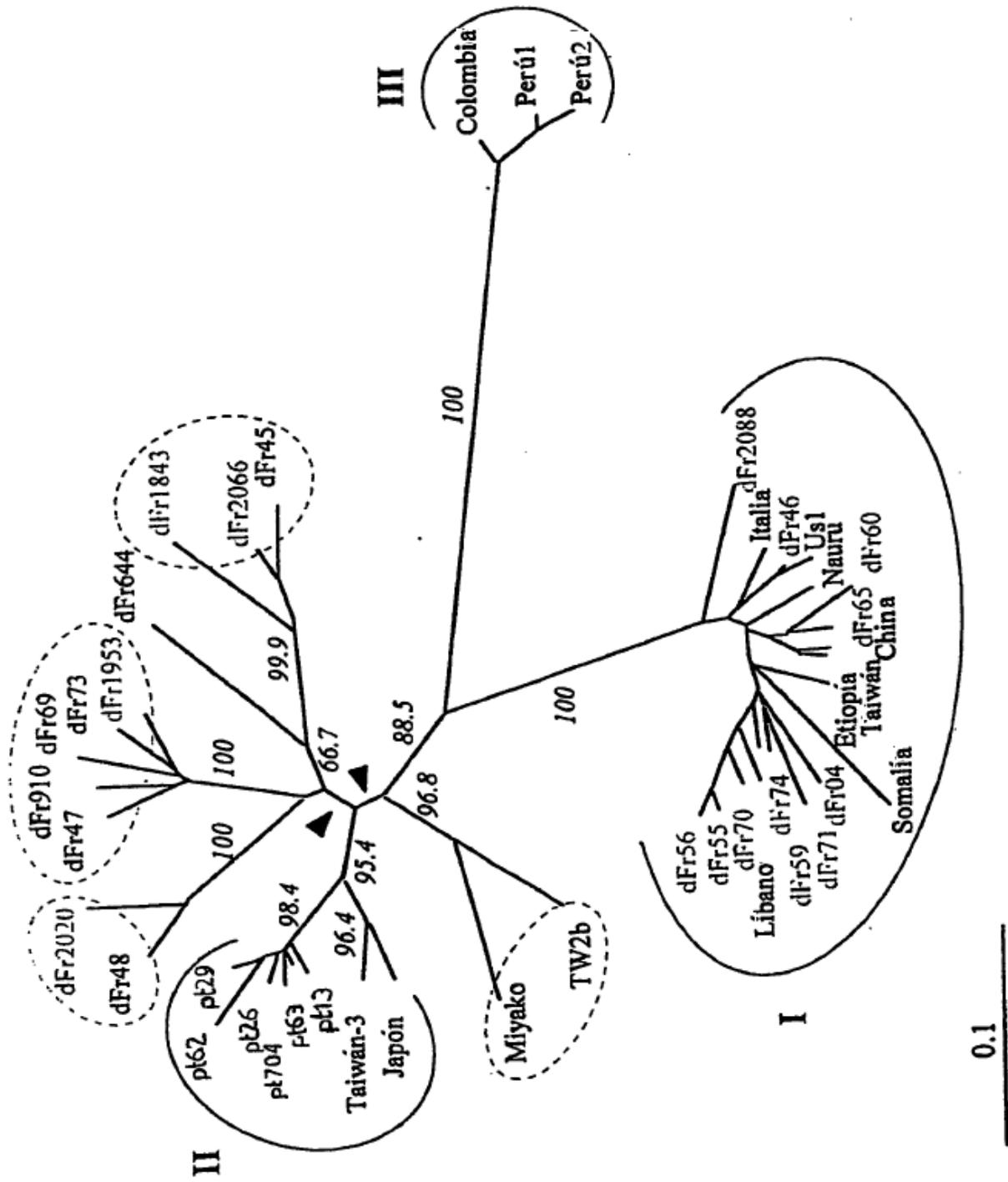


FIGURA 1

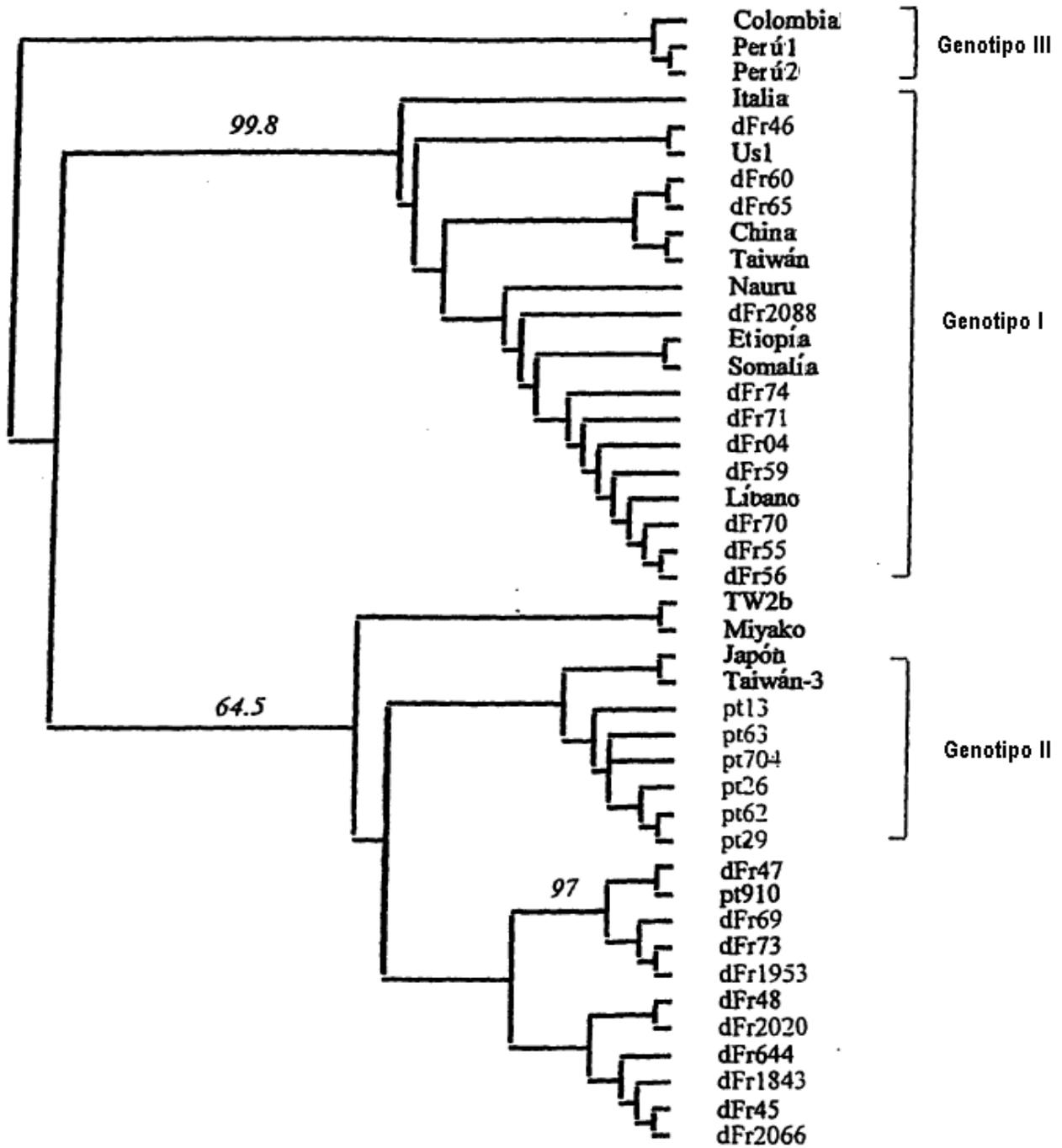


FIGURA 2

ES 2 387 604 T3

Pacientes				Marcadores de HDV				Hígado	Tratamiento
Nº de anonimato/ paciente	Sexo	Edad	Origen	Anti-HD			ARN	Anatomo- patología	Médico quirúrgico
				Ag- HD	IgM	IgG	PCR1/2*		
dFr45	F	43	Camerún	-	+	+	-/+	HCA	IFN x 3 9 M 12 meses
dFr47	F	38	Guinea	-	+	+	-/+	Cirrosis	NT
dFr48	F	40	Polonia (estancia en Camerún)	-	+	+	-/+	HCA	IFN x 3 5 M 6 meses
dFr73	M	32	Costa de Marfil	-	+	+	-/+	HCA	IFN x 3 10 M 12 meses
dFr644	F	30	Congo	-	+	+	-/+	Cirrosis	Transplante 9/00
dFr910	F	32	Malí	-	+	+	-/+	HCA	IFN x 3 3 M 3 meses

FIGURA 3

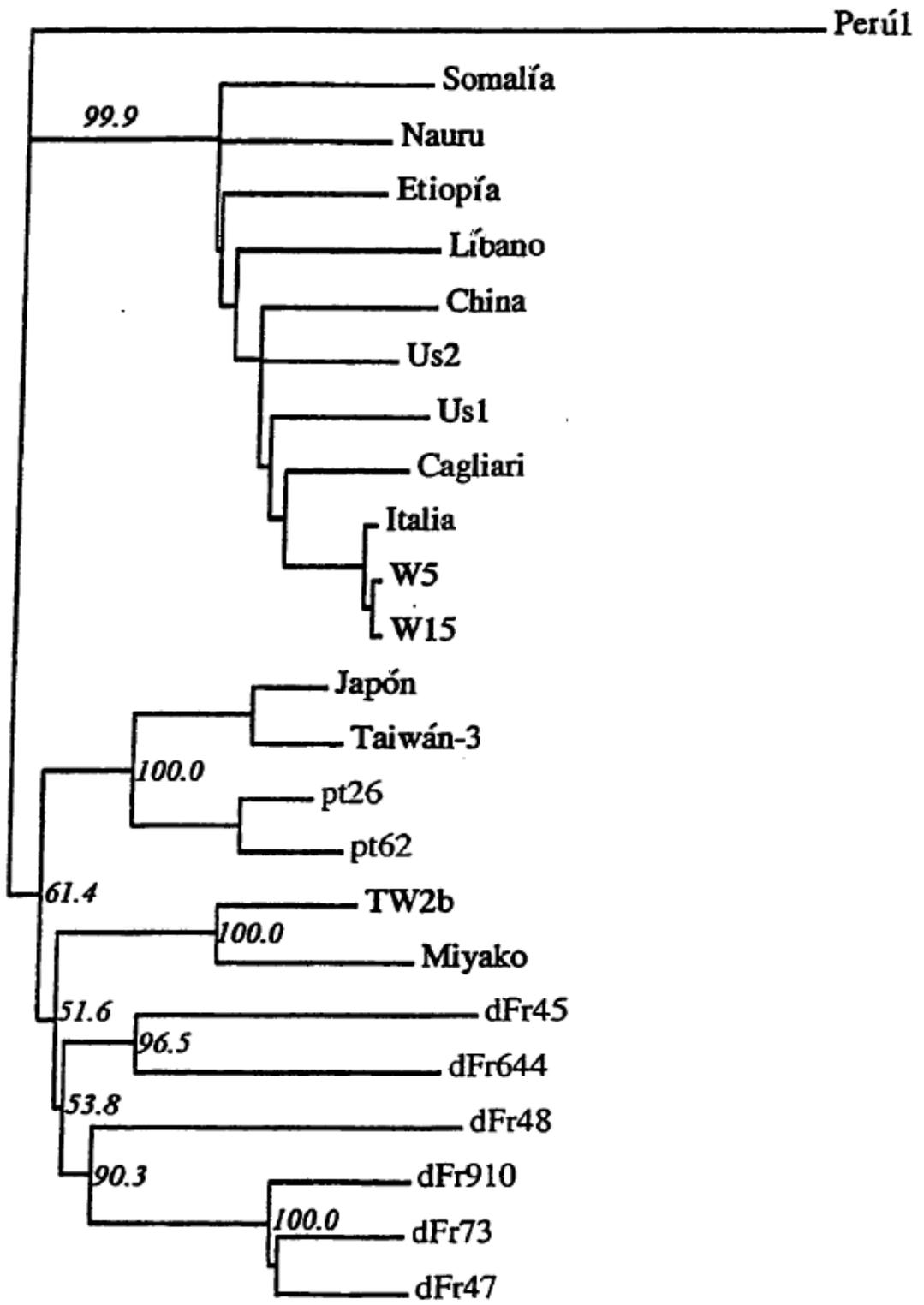


FIGURA 4

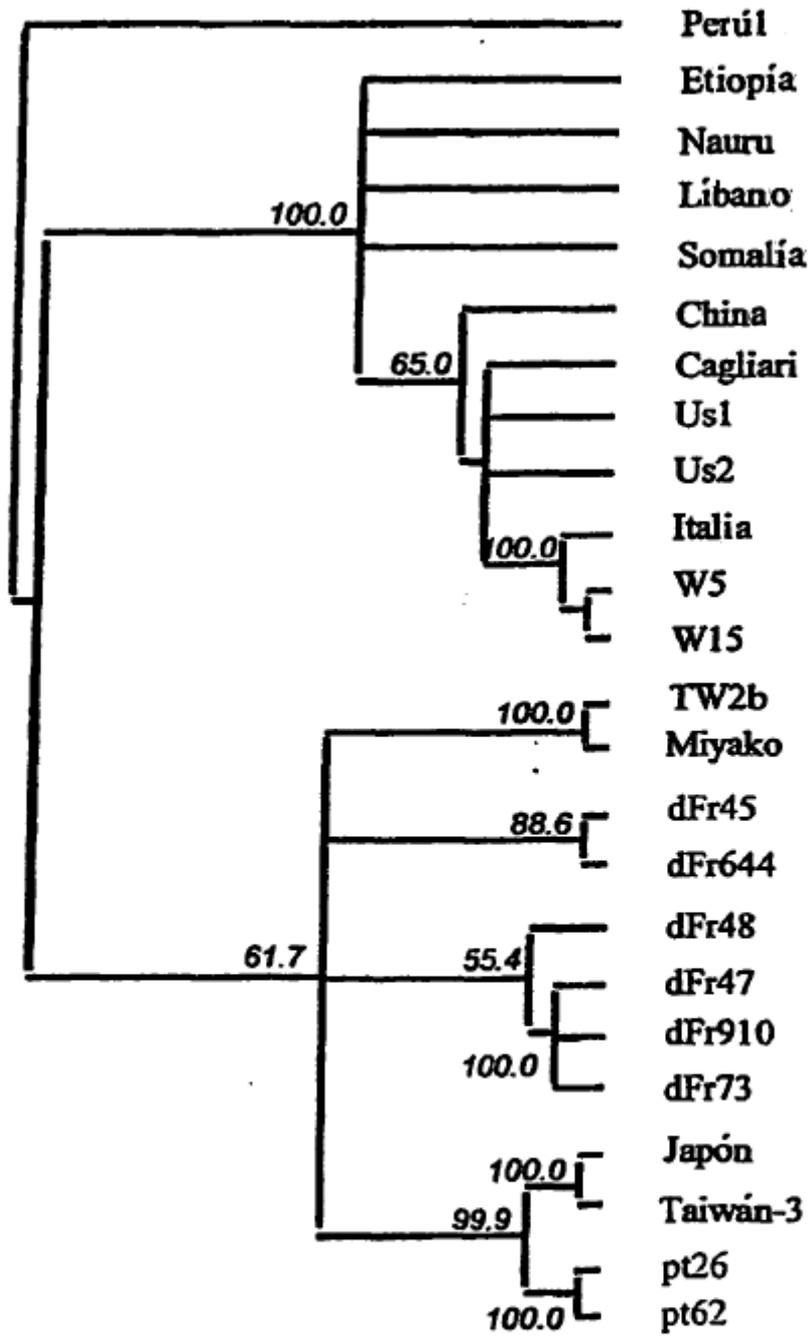


FIGURA 5

		1		50
1	TW2b	MSQPDSRRPRRG-REEQLGKWIDARRRKEELERDLRKNVNTIKRLEEDNPWLGNVIRGIIRK-DKD		
2	Miyako	...S.A.V...-...T...G...KE...I.NL...I...-...		
3	Taiwán-3	...SE..KS...T...T.ER..TT..KA...K...AR...K...E...IL...-G..		
4	Japón	...SET..G...T...T.E...T...A...K...TR...K...E...IV...-G..		
5	pt62	...SE..KS...G..DI.E..VST.KKA...AR...D...IL...-G..		
6	pt26	...E..KG...G..I.E..ST..KA...AR...DE...IL...-G..		
7	dFr910	...SE.KKA...G...I.E..VQ..KDS.D..KK...TKR...K..DE...IL...-G..		
8	dFr47	...SEQ.K...G..DV.E...Q..KDA.D..KR...AKR.A.K..DE...IL...-G..		
9	dFr73	...SEQKKS...G...T.E...Q..KDA.D..KR...TKR...DE...L...-G..		
10	dFr644	...S.AK.E...G..DV.S..VE..KDL.D..KRI..TRRN...DE...IL...-G.A		
11	dFr45	..HA.TK.S.K.-...T.S..DK..EDA...R...TK...K..D...IK..G.-VGT		
12	dFr48	.GPAEQK.-K..G...I.E..VEL.KNR.D...TQ.GL.K..D...II...-G..		
13	Etiopía	..RTE.KK-N..G..DA.EQ.VNG...I...D..AK.K..K..D...IK..LG.K..		
14	Italia	..RSE.RK-N..G...I.EQ.VAG.KKL...D..TK.KL.KI.DE...IK..LG.K..		
15	Somalia	..RSEQK-N..G..DT.E..VNE.KKA...KE...AK.K..K..E...IK..LG.K..		
16	Perú	...TVA.LTSKE-...I.EQ.VEE.KNRRK..KD..RA..K..K..DE...V.LL.R-K..		
		*. . . : ** : * : * * . . ** : * : . : * : * : * : * : * : * : * : * : *		
		100		
1	GEGAPPAKRARTDQMEVDSGPRKRKHGGFTEQERRDHRRRKALENKKKQLSSGGKNLSREEEEELRRLTEEDERR			
2V.....K.....:			
3S.....TG..P.RS...DK..E.....R.....S.....V..E.			
4P.....G..P.KS...DK..E.....A..I..K.....D..E.			
5KP.KS...DK..Q.....Q..N..A..S..K.....I..DE.			
6M..P.KS...DK..E.....R.....A.....			
7TRA.D..D..K.....N..A..S.....K.....			
8TRA.D..DK.....A..S..K.....K.....			
9R.....SRA.D..DK..Q.....NA..S..K.....G.....			
10G.KSRK...DE.....K..E..G.....V..K.			
11P..R.....G.KSNK...DE..A.....A..S..K.....K..AD..Q.			
12T.K.....GKP.KS...DE.....AA.....G..G..Q.			
13PLR...DK..Q.....A..H.....K.....			
14PLR...DK..Q.....A..K.....			
15P.....I.....KPLR...DK..E.P.....A.....K.....AR..E.			
16	ED.....P.QET.....GRKPKAR...D.....AG...H..Q.....ARD.DE.			
		:****:*:* : **:* : * : * : * : * : * : * : * : * : *		
		196		215
1	ERRVAGPRVGDVNPLDGGPRGAPGGGFVPSMHDIPESPFTRRGDGLDVRGAQEFZVSPQPPPRLELLE-CTPQ			
2E.A.....T.EGV...S.....I..T.G..ZVD.GR.S.....-...			
3	K.....V..SG.....Q.EGV...A.T.E..I..N.G..WVH.S..QQ.....-...			
4	K.....SR.....Q.AGV...S.T.E..I..T.G..ZV..S..QQ.....-...			
5	Q.....PG.S.....Q.TGV...S.T.E..I..DRG..WVN.A..GQ.....-...			
6PG.S.....Q.TGV...T.E..I..T.G..WVN.V..GQ.....-...			
7PE.....Q.L.V...S.T.S...N.Q..WVD.G.R.....-...			
8PGTS.....Q.LGV...S.T.E...N.QY.WVN.G.R.....-...			
9PQ.....Q.LNV...T.E..I..N.Q..WV..GA.S.....-...			
10	...A..QD.G..PG.S.....R.LGV.....I..D.Q..WGQRP.....-...			
11	A..I.....PE.....QLLGV...S.T.E..I..DRQ..WGSPPT.....-...			
12	K..T.....PG.D.....T.LHV...S.T.E...T.Q..WGNTP.R.....-...			
13	...A..P..G..E..Q.....QGV...H.E...TGG..ZDILF.SD.PFSPQS-.R..			
14	...P..G..I..E..S.....LQGV...S.T.E..I..NRG..WDILF.AD.PFSPQS-.R..			
15	S..I..SA.G..E..S.....QG...H..E...T.TGG..ZDILF.SD.PFSPQS-.R..			
16	...T...P.G..M..P.....LQGV...S.T.E.I.I..T.Q..WYGF...GYWVPG.TQ.			
		** *** * * * ***** : ***** * * : * : * : *		

FIGURA 6