



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 350 179**

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: **02781380 .7**

96 Fecha de presentación : **23.09.2002**

97 Número de publicación de la solicitud: **1430134**

97 Fecha de publicación de la solicitud: **23.06.2004**

54

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

30

Prioridad: **24.09.2001 FR 01 12285**

45

Fecha de publicación de la mención BOPI:
19.01.2011

45

Fecha de la publicación del folleto de la patente:
19.01.2011

73

Titular/es: **Assistance Publique, Hopitaux de Paris
3, avenue Victoria
75004 Paris, FR**

72

Inventor/es: **Deny, Paul;
Huc-Anais, Patricia y
Radjef, Nadja**

74

Agente: **Martín Santos, Victoria Sofía**

ES 2 350 179 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).

MOLÉCULAS DE ÁCIDOS NUCLEICOS DEL HDV, SUS FRAGMENTOS Y SUS APLICACIONES

DESCRIPCIÓN

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

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

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

20 Dentro de los viriones, el ARN del HDV es circular y de polaridad negativa. Esta cadena monocatenaria circular cerrada, el genoma más pequeño conocido de los virus que infectan a mamíferos, tiene un porcentaje elevado de GC (60%).

El ARN del HDV se replica independientemente del HBV, cuya función se limita a proporcionar la envoltura del HDV. Las únicas proteínas encontradas (sHD y LHD) las codifica el ARN antigenómico que, en la célula infectada, es completo, circular y pseudo-bicatenario, sirve de intermediario de la replicación y sostiene la edición.

25 El ARN del HDV pertenece a un tipo de ribozima específica. La reacción de auto-escisión necesita el ARN y un catión divalente (Mg⁺⁺). La escisión crea un extremo 2', 3' fosfato cíclico y un extremo 5' hidroxilo.

30 Las ribozimas delta (genómica y anti-genómica) tienen una configuración secundaria de similar a pseudo-nudos. Las secuencias implicadas incluyen principal o exclusivamente las secuencias situadas en 3' del sitio de auto-escisión (aproximadamente 84 nucleótidos).

Durante el ciclo viral, el ARNm del HDV codifica una proteína que existe en dos formas: una proteína de 194-195 aminoácidos (forma 's' que indica pequeña (small)) de 240 kilodaltons (kDa) y una proteína de 214 aminoácidos (forma 'L' que indica grande (Large)) de 27 kDa, que existen en proporciones variables. Estas proteínas llevan la antigenicidad 'delta' y se detectan en el hígado o en el suero de pacientes o animales infectados (chimpancé, marmota). Estas dos proteínas virales sHD y LHD comienzan en el primer ATG de la fase de lectura abierta situada en la posición 1598 (según la numeración de Wang et al., 1986 o 1987) del ARN antigenómico. Durante la replicación, en la posición

1012 aparece una mutación, dependiente de una enzima celular, la adenosina desaminasa, dependiente del ARN bicatenario, transformando el codón de terminación ámbar (UAG) en el codón del triptófano (UGG), alargando en dirección 3' la fase de lectura de 19 ó 20 codones y confiriendo a las dos formas sHD y LHD diferentes propiedades.

- 5 El ARNm acaba en una cola de poli(A), 15 nucleótidos después de la señal de poliadenilación consenso, AAUAAA (posiciones 954-959).

En el ciclo de replicación, las funciones de las proteínas de 24 y 27 Kd se oponen: sHD activa la replicación viral mientras que LHD la reprime y desempeña una función en el ensamblaje de las partículas virales. Estas proteínas se fosforilan en los restos de serina pero no se glicosilan (Tabla I).

- 10 Están constituidas por dominios funcionales comunes y por un dominio específico en la proteína grande LHD.

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	+	-
Trans-inhibición de la replicación	-	+
Dimerización y polimerización	+	+
Fijación al ARN	+	+
Estabilización del ARN	+	+
Localización nuclear	+	+
Ensamblaje	-	+
Fosforilación	+	+(x6)
19 aa carboxi-terminales específicos	-	+
Farnesilación	-	+

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

* dominios comunes

- 15 - El dominio de polimerización, que comprende la secuencia entre los restos aminoacídicos 13 y 48 formada por una disposición de leucina o isoleucina, organizada en hélice α de tipo “cremallera de leucina” (leucine zipper), implicada en la polimerización de las proteínas, indispensable para (i) la transactivación de la replicación viral por el Ag-sHD, (ii) la inhibición de la replicación por el Ag-LHD y (iii) el ensamblaje de los complejos sHD-LHD en las envolturas del HBV.
- 20 - La señal de localización nuclear (SLN) que implica dos secuencias de localización nuclear identificadas en la región 67-68, esenciales para translocar el Ag-sHD sintetizado el citoplasma y quizá la ribonucleoproteína después de su entrada en la célula, hacia el núcleo.
- El sitio de fijación al ARN que se basa en dos secuencias ricas en arginina localizadas entre los restos 97 y 163 que permiten la unión de las proteínas sHD sobre el ARN genómico
- 25 o antígenómico. Esta fijación es indispensable para que el Ag-sHD active la replicación.

* dominios específicos

Los 19-20 aminoácidos situados en el extremo COOH de la proteína grande tienen una función importante en el ciclo del HDV. En efecto, estos aminoácidos (aa 195-214), intervienen en el ensamblaje de las partículas virales (Chang et al., 1991). Esta actividad podría estar en parte relacionada con la presencia de una cisteína en la posición 211 (Glenn et al., 1992), conservada para todos los genomas virales caracterizados hasta ahora. Esta cisteína, situada 4 aminoácidos delante del extremo COOH de la proteína, forma una caja “CXXX” y fija un grupo farnesilo (Glenn et al., 1992) cadena de 15 carbonos derivada del ácido mevalónico, por acción de una farnesiltransferasa. Esta maduración post-traducciona orienta las proteínas hacia las membranas celulares.

Por otro lado, la proteína pequeña y grande se han diferenciado por anticuerpos monoclonales (clon 9E4) (Hwang et Lai, 1993a). Estos anticuerpos sólo reconocen la sHD (Lai et al., 1993). Como la secuencia de aminoácidos de la proteína pequeña se incluye en la grande, estos resultados sugieren una conformación diferente entre sHD y LHD en el seno de los 30 aminoácidos carboxiterminales de la proteína pequeña sHD, lo que sugiere que el epítipo reconocido sobre la sHD debe ocultarse en la LHD en condiciones no desnaturalizantes.

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

En América del Norte y en Europa del Oeste, la hepatitis D se localiza sobre todo en consumidores de drogas intravenosas, hemofílicos y los que reciben múltiples transfusiones.

La epidemiología y los medios de contaminación se superponen en parte. Se estima globalmente en un 5% la proporción de portadores del Ag-HB infectados por el HDV. Por lo tanto, se han constatado las disparidades de prevalencia geográfica y epidemiológica.

En determinadas regiones del globo, existe una alta prevalencia de esta enfermedad, en los portadores del virus de la hepatitis B, incluyendo la cuenca amazónica de América del Sur, África central, sur de Italia y los países del Oriente Medio.

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

En África y en Asia, donde el portador crónico del HBV es de frecuencia elevada (del 10% al 20%), así como en América del Sur y en las islas del Pacífico, donde es intermedia (del 1% al 5%), la distribución del HDV es paradójicamente disparatada. En África, los estudios de seroprevalencia muestran una distribución muy heterogénea de pacientes que tienen anticuerpos anti-HD, mientras que la prevalencia global de la infección por HBV, calculada para la detección del Ag-HBs, se estabiliza entre el 12 y el 14% (Roingeard et al., 1992). De esta manera, tasas variables del 4% (zona Norte de Senegal) al

44% (suburbios de Dakar) hacen aflorar probables factores socioeconómicos implicados en la transmisión.

Los estudios de prevalencia del HDV deben interpretarse prudentemente. De hecho, en las poblaciones estudiadas, existe una inclusión preferencial de pacientes que padecen hepatopatías. En los pacientes que padecen hepatitis aguda o crónica, la prevalencia de la infección por el HDV es más importante que en los portadores crónicos asintomáticos del HBV. Además, la investigación serológica de una infección HDV se basa en la detección del Ag-HD y de los Ac anti-HD totales en el suero. Debido a esto, infecciones agudas benignas, en el transcurso de las cuales se desarrollaría una producción transitoria aislada de IgM anti-HD, no se volverían a censar.

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. Es una de las razones por las que el diagnóstico fiable del HDV asociado a HBV es crucial.

La infección por un HDV es dependiente del HBV. Los aislados del HDV de regiones geográficas diferentes muestran una variabilidad genética. Actualmente se identifican tres genotipos y se denominan genotipo I, II y III.

El genotipo se utiliza en la epidemiología de las transmisiones virales, permite estudiar la distribución geográfica y podría correlacionarse con el poder patogénico.

El HDV sólo se desarrolla en los pacientes también infectados por el HBV. Esta doble infección deriva de una co-infección o de una sobreinfección:

- la co-infección es la causante de una hepatitis aguda. El diagnóstico, evocado en el transcurso de una citólisis hepática, se basa en la detección de marcadores del HDV asociada 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 constatar una inhibición de replicación del HBV por el HDV. Las IgM anti-HBc demuestran la infección reciente por el 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: las IgM anti-HD predominan, pero guardan una titulación moderada ($<1:1000$). Se observan dos picos de elevación de transaminasas, separados de dos a cinco semanas, en el 10 al 20% de las co-infecciones, reflejando probablemente diferentes cinéticas de replicación de virus. La co-infección se caracteriza por tanto por una gravedad de la hepatitis aguda a menudo superior a la ocasionada por el HBV aislado. De esta manera, se describen hepatitis fulminantes en América del Sur y en África subsahariana o en determinadas poblaciones. La evolución está generalmente marcada por una resolución de la hepatitis después de la fase aguda y a la vista de la historia natural del HBV, sólo el 5% de los pacientes co-infectados evolucionan hacia la cronicidad.

- La sobreinfección se caracteriza por la aparición de una seroconversión HDV en el paciente portador crónico del Ag-HBs. La viremia HDV precede la aparición de anticuerpos anti-HDV en ausencia de detección de las IgM anti-HBc. La detección de estos marcadores puede preceder una elevación de transaminasas de varios meses. En fase aguda, la sobreinfección se traduce por una hepatitis

- fulminante en más del 10% de los casos. Además, pasada la fase aguda, la sobreinfección arrastra frecuentemente (del 60 al 70%) una hepatitis crónica activa con evolución rápida hacia la cirrosis. En la fase aguda de la sobreinfección, la detección del Ag-HD se produce rápidamente después de la aparición de anticuerpos que persisten a tasas elevadas. Contrariamente a los modelos clásicos de infección viral, la
- 5 IgG anti-HD y la IgM anti-HD se detectan simultáneamente en las hepatitis B-delta crónicas.

La siguiente Tabla II resume la evolución de marcadores B y delta durante co-infecciones y sobreinfecciones.

Co-infección por el HDV	Fase aguda	Evolución	
		Crónica	Curación
Ag-HBs	+	+	-
IgM anti-HBc	+	-	-
Ag-HD	+/-	-	-
IgM anti-HD	+/-	+	-
IgG anti-HD	+/-	+	+
ARN HDV	+	+	-
Ag-HD intrahepática	+	+	-
Sobreinfección por el HDV		Evolución	
		Crónica	Curación
Ag-HBs	+	+	-
IgM anti-HBc	-	-	-
Ag-HD	+/-	-	-
IgM anti-HD	+	+	-
IgG anti-HD	+	+	+
ARN HDV	+	+	-
Ag-HD intrahepática	+	+	-

- La co-infección y la sobreinfección son clínicamente indistintas. El diagnóstico virológico se basa habitualmente en diferentes marcadores séricos. Más raramente, el Ag-HD puede detectarse en cortes anatomopatológicos de la biopsia de hígado.
- 10

Los marcadores permiten seguir la evolución de la enfermedad hacia la curación o la cronicidad, decidir el tratamiento de una enfermedad o evaluar la eficacia.

- El HDV sólo se aísla en cultivos celulares si el diagnóstico se basa esencialmente en la investigación del Ag-HD (ELISA, IF) o del genoma viral (hibridación, PCR, PCR en tiempo real) para las técnicas directas y en la detección de anticuerpos IgM anti-HD e IgG anti-HD para los métodos indirectos (ELISA).
- 15

- La investigación del Ag-HD intrahepático puede discutirse en las hepatitis fulminantes teniendo en cuenta la cinética de aparición de los marcadores séricos. Este examen tiene un valor de referencia para estudiar la replicación del HDV, pero no puede utilizarse rutinariamente.

- El Ag-HD sérico se investiga en el suero en presencia de un agente disociante que expone el Ag-HD, incluido en la envoltura vial portando el Ag-HBs. La presencia de anticuerpos (Ac) anti-HD a índices elevados (hepatitis crónicas) que fijan los antígenos séricos impide la detección. Las técnicas de transferencia de Western se desarrollan con objeto de investigación. La presencia del virus en la sangre es transitoria y limitada en la fase precoz de la infección y la posibilidad de detectar el Ag-HD disminuye días después de la aparición de los síntomas.

- La inmunocaptura se utiliza para detectar los IgM anti-HD y la competición para los IgG anti-HD. Las técnicas ELISA han utilizado en primer lugar como antígeno el Ag-HD de suero o de hígado de pacientes o de animales infectados. Los nuevos ensayos se basan en Ag-HD recombinantes o en péptidos de síntesis.

- Las técnicas de hibridación o de RT-PCR permiten la detección del ARN genómico tras la extracción de ácidos nucleicos y la desnaturalización de estructuras secundarias. Se han descrito varios sistemas de cebadores: su elección es determinante ya que la variabilidad genética en las regiones denominadas conservadas puede conducir a falsos negativos si los cebadores seleccionados no son los apropiados a las cepas virales circulantes. 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.

Sin embargo, también tanto en caso de co-infección como en caso de sobreinfección, el Ag-HD es en efecto difícil de detectar, ya que la viremia precede la aparición de anticuerpos.

En este contexto, y particularmente gracias a la puesta en evidencia de nuevos genotipos, son necesarios reactivos nucleicos y proteicos que permitan el diagnóstico del HDV, cualquiera que sea el genotipo.

En efecto el estudio de las secuencias de nucleótidos del HDV, por diferentes equipos en el mundo, ha permitido diferenciar, hasta ahora, sólo tres genotipos distintos:

- el genotipo-I, 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 del HDV de aislados geográficos diferentes. La primera secuencia de un HDV en el ser humano 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, en los Estados Unidos, Taiwán, Nauru, Francia, Líbano, China (Makino et al., 1987; Chao et al., 1991b; Imazeki et al., 1991; Lee et al., 1992; Niro et al., 1997; et Shakil et al., 1997). En el interior del genotipo I se describe un porcentaje de similitud de nucleótidos de más del 85%.

- Un aislado japonés (Imazeki et al., 1990; Imazeki et al., 1991) es el prototipo de un 2º subgrupo de HDV. Este genotipo II que no se había descrito en un primer momento más que en Japón y Taiwán (Imazeki et al., citados anteriormente; Lee et al., 1996b) parece tener una distribución geográfica mucho más grande. En particular, también se han caracterizado secuencias del genotipo I y del genotipo II provenientes de Yakouti (Rusia). Finalmente, algunos autores se basan en una diversidad

5 - En 1993 se describió un 3^{er} grupo para genomas de virus peruanos y colombianos
(Casey et al., 1993a). El genotipo-III sólo se ha descrito en América del Sur, y más particularmente en la
cuenca Amazónica asociado a hepatitis graves, es decir epidemias de hepatitis fulminantes con esteatosis
microvesicular (Casey et al., 1993a; Casey et al., 1996b) y con una morbilidad y una mortalidad elevadas.
En esta región geográfica, se observa que el HDV de genotipo-III está asociado de manera privilegiada
10 con el HBV de genotipo F. Otros aislados de este grupo se han aislado recientemente en Venezuela
(Nakano et al., 2001).

15 dos regiones conservadas se sitúan en la fase de lectura que codifica la proteína HD (Chao et al., 1991b).

20 En consecuencia, las técnicas de detección especificadas anteriormente corren el riesgo de llegar a resultados negativos tanto a nivel nucleico como lo que concierne a la respuesta de anticuerpos.

25 hibridación), lo suficientemente sensibles y específicos, es decir que no conduzcan a resultados falsamente negativos o falsamente positivos: de hecho, actualmente, en el contexto de una hepatitis grave, puede observarse una disociación IgM anti-HD positiva / ARN-HDV negativa.

30 anteriormente descrita, lo que tiene consecuencias sobre la fiabilidad del diagnóstico.

En particular han puesto en evidencia nueve secuencias completas nuevas del HDV (tres originarias de Yakouti y seis originarias de África), que también circulan en la isla de Francia y que no pertenecen a ninguno de los genotipos conocidos. El análisis de estos nuevos aislados:

- 35 que se había descrito hasta ahora,
- replantea la clasificación de los HDV de solamente tres genotipos
 - ha llevado a los inventores a proponer un algoritmo de PCR-RFLP, desde una
- región parcial de genoma para el genotipado HDV, y

- ha llevado a los inventores a desarrollar reactivos adaptados a un diagnóstico fiable de infecciones del HDV y cualquiera que sea el genotipo, mientras que anteriormente, se observaban numerosos resultados falsamente negativos (existencia de nuevos genotipos).

Los inventores se propusieron por tanto proporcionar moléculas de ácido nucleico de HDV, adecuadas para permitir la detección de una variante de HDV en relación con los tres genotipos anteriormente descritos.

Se describen moléculas de ácido nucleico aisladas, **caracterizadas por que se seleccionan del grupo constituido por:**

- el genoma de un HDV, que molecularmente presente, en su genoma completo, una divergencia o distancia genética $\geq 20\%$ (menos del 80% de similitud) con respecto a las secuencias de un HDV de genotipo I, un HDV de genotipo II o un HDV de genotipo III,
- el genoma de un HDV, que molecularmente presenta una divergencia o distancia genética $\geq 25\%$ (menos del 75% de similitud) sobre una región denominada R0, delimitada por las posiciones 889 a 1289 de genoma HDV, con respecto a las secuencias R0 correspondientes a un HDV de genotipo I, un HDV del genotipo II o un HDV de genotipo III,
- los genomas completos de aislados o variantes del HDV denominadas dFr45, dFr47, dFr73, dFr910, dFr48 y dFr644 que presentan la secuencias SEC ID N°: 1, 6, 11, 16, 21 y 26, respectivamente y
- el genoma de un HDV que presenta una divergencia o distancia genética $\leq 15\%$ con al menos una de las secuencias de las SEC ID N°: 1, 6, 11, 16, 21 y 26.

De acuerdo con un modo de realización ventajoso de dichas moléculas, la región R0 se obtiene preferentemente por amplificación del ARN del HDV con los cebadores 900S (SEC ID N°: 33) y 1280AS (SEC ID N°: 34).

En el contexto de la presente invención, se entiende por molécula de ácido nucleico, una molécula de ADNc o de ARN que presenta una de las secuencias genómicas del HDV tales como se definen anteriormente y sus secuencias sentido y antisentido complementarias.

También se describen moléculas de ácido nucleico que comprenden al menos uno de los fragmentos de las secuencias de una variante del HDV tales como se definen anteriormente, seleccionadas en el grupo constituido por:

- a) los fragmentos R0 de las siguientes variantes del HDV aisladas : dFr45, dFr47, dFr48, dFr69, dFr73, dFr644, dFr910, dFr1843, dFr1953, dFr2020 y dFr2066, que presentan las siguientes secuencias: SEC ID N°: 48 a SEC ID N°: 58, respectivamente,
- b) el fragmento R1 que se extiende desde las posiciones 307 a 1289 del genoma del HDV,
- c) el fragmento R2 que se extiende desde las posiciones 889 a 328 del genoma del HDV,
- d) el fragmento R3 que se extiende desde las posiciones 1486 a 452 del genoma del HDV,
- e) el fragmento R'1 que se extiende desde las posiciones 305 a 1161 del genoma del HDV,
- f) el fragmento R'2 que se extiende desde las posiciones 984 a 331 del genoma del HDV,

- g) el fragmento R644 que se extiende desde las posiciones 889 a 446 del genoma del HDV,
- h) el fragmento G910 que se extiende desde las posiciones 1206 a 929 del genoma del HDV,
- i) el fragmento p910 que se extiende desde las posiciones 553 a 1550 del genoma del HDV,
- j) los ADNc que codifican la proteína sHD, de las secuencias SEC ID N°: 4, 9, 14, 19, 24 y 29,
- 5 k) los ADNc que codifican la proteína LHD, de las secuencias SEC ID N°: 2, 7, 12, 17, 22 y 27 y
- l) los cebadores de las secuencias SEC ID N°: 33 a SEC ID N°: 47.

En el contexto de la presente invención, las posiciones de los fragmentos en el genoma del HDV son idénticas en el genoma circular de orientación genómica, de acuerdo con la numeración de Wang y al., 1986 ó 1987.

10 También se describen fragmentos de nucleótidos complementarios de los anteriores así como fragmentos modificados, con respecto a los anteriores, por supresión o adición de nucleótidos en una proporción de aproximadamente el 15%, con respecto a la longitud de los fragmentos anteriores y/o modificados a nivel de la naturaleza de nucleótidos ya que los fragmentos de nucleótidos modificados conservan una capacidad de hibridación con las secuencias genómicas o antigenómicas del

15 ARN de aislados tales como los definidos anteriormente.

De hecho, estas diferentes cepas virales, en un mismo paciente, en un momento determinado, muestran una población heterogénea de moléculas de ARN del HDV; por otro lado, durante una infección crónica, además de las heterogeneidades observadas a nivel del sitio de edición (posición 1012), pueden aparecer mutaciones. Las secuencias virales parecen evolucionar dentro de las poblaciones

20 virales con una tasa de sustitución variable de 3.10^{-2} a 3.10^{-3} por nucleótido y por año.

Algunos de esos fragmentos son específicos y se utilizan como sonda o como cebadores; se hibridan específicamente con una cepa de una variante del HDV tal como se define anteriormente o con una cepa emparentada; se entiende por HDV emparentada con una variante tal como se define anteriormente, un HDV que presenta una divergencia genética $\leq$ al 15%.

25 Dichos fragmentos se utilizan para la detección y el seguimiento epidemiológico de infecciones por el HDV. Por ejemplo, el fragmento R0 se utiliza para la detección (RT-PCR) y el genotipado (PCR-RFLP) del HDV. Los otros fragmentos que cubren la totalidad del genoma del HDV se utilizan para la caracterización molecular de las variantes del HDV; el análisis filogenético de la secuencia completa del genoma o fragmentos de éste que corresponden en particular a R0 o a R2

30 permiten relacionar los perfiles observados en PCR-RFLP con un genotipo proporcionado o caracterizar nuevos genotipos.

En consecuencia, también se describe un método de detección de una variante del HDV de acuerdo con la invención, por hibridación y/o amplificación, realizado a partir de una muestra biológica, en el que el método se **caracteriza por que** comprende:

- 35 (1) una etapa de extracción del ácido nucleico a detectar, perteneciente al genoma del virus eventualmente presente en la muestra biológica,

(2) la realización de al menos una amplificación génica usando un par de cebadores seleccionado del grupo constituido por los cebadores adecuados para amplificar una de las siguientes regiones del ARN genómico del HDV: R0, R1, R2, R3, R644, G910, p910, R'1 y R'2 y

(3) el análisis del producto amplificado por comparación con una de las moléculas de las secuencias SEC ID N°: 1, 6, 11, 16, 21 y 26, que corresponden a los genomas completos de aislados o variantes denominadas dFr45, dFr47, dFr73, dFr910, dFr48 y dFr644, respectivamente

De manera ventajosa, la etapa (3) del análisis puede realizarse por restricción, secuenciación o hibridación; en el este último caso, la sonda utilizada (particularmente en microplacas de ADN) será ventajosamente un fragmento de 15 a 20 nucleótidos, específico de dichos fragmentos amplificados.

De acuerdo con un modo de realización 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, realizadas en la etapa (2) se seleccionan en el grupo constituido por:

- los cebadores 900S (SEC ID N°: 33) y 1280AS (SEC ID N°: 34), para la amplificación de R0 (aproximadamente 400 pb),
- los cebadores 320S (SEC ID N°: 39) y 1280AS (SEC ID N°: 34), para la amplificación del fragmento R1 (aproximadamente 960 pb),
- los cebadores 900S (SEC ID N°: 33) y 320AS (SEC ID N°: 45), para la amplificación de R2 (aproximadamente 1100 pb), que contiene el gen sHD que corresponde a las posiciones 1598-950,
- los cebadores 1480S (SEC ID N°: 46) y 440AS (SEC ID N°: 47), para la amplificación de R3 (aproximadamente 650 pb),
- los cebadores 900S (SEC ID N°: 33) y 420AS (SEC ID N°: 40), para la amplificación de la región R644 (aproximadamente 1250 pb) del aislado dFr644,
- los cebadores 318S (SEC ID N°: 35) y 1150AS (SEC ID N°: 36), para la amplificación de R'1 (aproximadamente 850 pb),
- los cebadores 960S (SEC ID N°: 37) y 345AS (SEC ID N°: 38), para la amplificación de R'2 (aproximadamente 1050 pb),
- los cebadores R910S (SEC ID N°: 41) y R910As (SEC ID N°: 42) para la amplificación de la región G910 (aproximadamente 1400 pb) del aislado dFr910,
- los cebadores S910R (SEC ID N°: 43) y As1910R (SEC ID N°: 44) para la amplificación de la región p910 (aproximadamente 650 pb) del aislado dFr910.

También se describe un método de detección y de genotipado del HDV a partir de una muestra biológica, cuyo método se **caracteriza por que** comprende:

- (a) una etapa de extracción del ácido nucleico perteneciente al genoma del virus HDV,
- (b) una etapa de amplificación de la región R0 delimitada por las posiciones 889 a 1289 del genoma del HDV,

(c) un primer tratamiento de las moléculas de ácidos nucleicos amplificados por las enzimas de restricción *SmaI* y *XhoI*, para producir un primer conjunto de fragmentos de restricción, y

(d) un segundo tratamiento de las moléculas de ácidos nucleicos por la enzima de restricción *SacII*, para producir un segundo conjunto de fragmentos de restricción

5 (e) el análisis combinado de dos conjuntos de fragmentos de restricción producidos por RFLP (Polimorfismo de longitud de fragmentos de restricción (*Restriction Fragment Length Polymorphism*)), para detectar la presencia y/o determinar el tipo de HDV presente en dicha muestra biológica.

De acuerdo con un modo de realización ventajosa de dicho método, la etapa (b) de
10 amplificación se realiza ventajosamente con los cebadores 900S (SEC ID N°: 33) y 1280AS (SEC ID N°: 34).

El método de acuerdo con la invención permite definir nuevos perfiles de restricción y clasificar los HDV en siete genotipos distintos.

De acuerdo con otro modo de realización ventajoso de dicho método, este comprende
15 además:

(f) la amplificación de las moléculas de ácido nucleico de dicha muestra por RT-PCR con los cebadores 900S (SEC ID N°: 33) y 320AS (SEC ID N°: 45) para amplificar la región R2 y

(g) la secuenciación directa de la región R2 amplificada y la comparación con una de las moléculas de ARN de las secuencias de las SEC ID N°: 1, 6, 11, 16, 21 y 26, que corresponden respectivamente
20 a los genomas completos de los aislados o de las variantes denominadas dFr45, dFr47, dFr73, dFr910, dFr48 y dFr644 respectivamente, por ejemplo, por análisis filogenético.

Cuando se observan perfiles no habituales, esta etapa complementaria permite caracterizar nuevos genotipos. De hecho, estos análisis complementarios de la PCR-RFLP permiten relacionar el nuevo perfil observado con un genotipo proporcionado, o caracterizar un nuevo genotipo,
25 por análisis Filogenético.

También se describe un vector recombinante, particularmente un plásmido que comprende un inserto constituido por una molécula de ácido nucleico tal como se define anteriormente.

También se describe una célula transformada por una molécula de ácido nucleico tal como se define anteriormente.

30 Se describen también productos de traducción codificados por una de las moléculas de ARN de las secuencias SEC ID N°: 1, 6, 11, 16, 21 y 26 que corresponden respectivamente a los ARN genómicos completos de aislados o variantes denominadas respectivamente dFr45, dFr47, dFr73, dFr910, dFr48 y dFr644 o por sus secuencias sentido o antisentido complementarias

También se describen las proteínas codificadas por el genoma de una variante del HDV
35 tal como se define anteriormente.

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

- la proteína LHD de dFr45, dFr47, dFr73, dFr910, dFr48 y dFr644 que presenta respectivamente las secuencias SEC ID N°: 3, 8, 13, 18, 23 y 28 y
- la proteína sHD de dFr45, dFr47, dFr73, dFr910, dFr48 y dFr644 que presenta respectivamente las secuencias SEC ID N°: 5, 10, 15, 20, 25 y 30.

5 La presente invención también tiene por objeto un péptido, **caracterizado por que** 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 carboxi-terminal de las secuencias SEC ID N°: 3, 8, 13, 18, 23 y 28,
- 10 - el péptido B de secuencia (código de una letra) RLPLLECTPQ (SEC ID N°: 59) constituido por los 10 aminoácidos del extremo carboxi-terminal de las secuencias SEC ID N°: 3, 8, 13, 18, 23 y 28, y
- el péptido C constituido por los 9 aminoácidos anteriores a la secuencia SEC ID N°: 59 (SEC ID N°: 60 a SEC ID N°: 65).

15 Dichos péptidos se utilizan para el diagnóstico indirecto (serología) de una infección por HDV, particularmente para un método inmunoenzimático (ELISA):

- el péptido B, que está conservado, permite la detección del conjunto de las variantes de acuerdo con la invención y del genotipo II del HDV, y
- el péptido C es específico de diferentes variantes del HDV de acuerdo con la invención.

20 Para la preparación de un kit de detección y del genotipado de un HDV, se describe la utilización de una molécula de ácido nucleico tal como se define anteriormente o de una proteína tal como se define anteriormente.

Además de las disposiciones anteriores, también se describen otras disposiciones, que se desprenderán de la siguiente descripción, que se refiere a ejemplos de realización de la presente invención así como a los dibujos adjuntos en los que:

- 25 - la figura 1 representa el árbol filogenético de la región R0 obtenido por el método del vecino más próximo “*neighbor-joining*”. Los números en cursiva indican los valores de *bootstrapping* (BV) (valores de remuestreo) sobre 10^4 remuestreos y el signo π indica los BV < 50%. La escala representa el número de sustitución de nucleótidos por sitio.

30 - la figura 2 representa el árbol filogenético de las regiones R0 del HDV obtenido por el método de máxima parsimonia. Los aislados Perú1, Perú2 y Colombia se seleccionan como “outgroup” (“grupo externo”). Las cifras en cursiva indican los valores de remuestreo (BV) sobre 10^4 remuestreos,

- la figura 3 ilustra los datos clínicos de cada uno de los seis pacientes infectados por los aislados del HDV de origen africano.

35 * indica las PCR 6S/6As y PCR R0, respectivamente

- la figura 4 representa el árbol filogenético de genomas completos del HDV, obtenido por el método del vecino más próximo “*neighbor-joining*”. Los números en cursiva, a nivel de

cada nudo, indican los valores de remuestreo (BV) sobre 10^4 remuestreos. La escala representa el número de sustitución de nucleótidos por sitio.

- la figura 5 representa el árbol filogenético de genomas completos del HDV, obtenido por el método de máxima parsimonia. Los números en cursiva, a nivel de cada nudo, indican los valores de remuestreo (BV) sobre 10^4 remuestreos.

- la figura 6 representa el alineamiento de las secuencias de aminoácidos de las proteínas delta de seis aislados de origen africano (líneas 7, 8, 9, 10, 11 y 12) con las secuencias conocidas del genotipo I (líneas 13, 14 y 15), genotipo II (líneas 3, 4, 5 y 6), genotipo III (línea 16) y TW2b/Miyako (líneas 1 y 2), usando el programa informático Clustal (versión 1.8). El aminoácido en la posición 196 de la proteína p27 corresponde al codón de terminación de la proteína p24 (Z) o al codón triptófano (W) que conduce a la síntesis de la proteína p27 que se extiende desde los aminoácidos 1 a 215.

Debe entenderse, sin embargo, que estos ejemplos se proporcionan únicamente a modo de ilustración del objeto de la invención, por lo que no constituyen de modo alguno ninguna limitación.

EJEMPLO 1: MATERIALES Y MÉTODOS

1- Pacientes y muestras

Se analizaron 22 sueros procedentes de sujetos sometidos a seguimiento en diferentes centros hospitalarios de la región parisina. Los pacientes eran portadores crónicos del Ag-HBs. El diagnóstico de la infección delta se efectuó investigando marcadores serológicos (Ag-HD, anti-AgHD de tipo IgM e IgG) y la detección del genoma viral del HDV por RT-PCR. En ninguno de los sueros analizados se detectó Ag-HD. En todos los pacientes se encontraron anticuerpos anti-AgHD del tipo IgM que manifiestan la cronicidad de la infección delta y de tipo IgG. En seis pacientes se caracterizó todo el genoma del HDV. Hasta la extracción del ARN viral, todos los sueros se conservaron a -80°C .

2- Extracción del ARN del HDV

Para extraer el ARN del HDV, se añadió un volumen de 250 μl de suero a 75 μl de Reactivo TRIzol RS (Gibco BRL, Life Technologies). Después de 30 segundos de homogeneización, la mezcla se incubó durante 5 min a temperatura ambiente. La extracción de los lípidos se efectuó añadiendo 200 μl de cloroformo enfriado a $+4^{\circ}\text{C}$. Después de otra homogeneización sometida a agitación vorticial, los tubos se incubaron y después se centrifugaron a 14000 revoluciones por minuto (rpm) durante 10 min a $+4^{\circ}\text{C}$. La fase acuosa se transfirió a tubos de extracción y los ARN se precipitaron por 500 μl de isopropanol frío, en presencia de 1 μg de glucógeno. Después de 15 min de homogeneización, las muestras se centrifugaron a 14000 rpm durante 10 min, a $+4^{\circ}\text{C}$. Después de aclarar con etanol al 70%, los tubos se centrifugaron de nuevo a 14000 rpm durante 10 min, a $+4^{\circ}\text{C}$. Los sedimentos se secaron en una campana a temperatura ambiente, después se volvieron a llevar a 100 μl de agua estéril que comprendía un inhibidor de ribonucleasas (RNasin, Promega). En esta etapa, se tomaron precauciones para evitar eventuales contaminaciones por parte de las ribonucleasas en los tampones y en las muestras.

3- Síntesis de un ADN complementario (ADNc)

Esta etapa consiste en sintetizar una cadena de ADN complementaria del ARN del HDV por transcripción inversa.

Para eliminar la estructura secundaria del ARN del HDV, se añadieron 5 µl de ARN del extracto anterior a una mezcla de reacción que contenía 5 µl (ó 0,5 pmoles) de desoxinucleótidos trifosfato (dNTP) y 1 µl (0,4 pmoles) de hexanucleótidos aleatorios. A continuación los ARN se desnaturalizan durante 3 min a 95°C. Para fijar los ARN desnaturalizados, los tubos se congelan inmediatamente en etanol enfriado a -20°C. Diez µl de una mezcla de reacción, que contenía 2,5 µl de ditiotreitól (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) se añadieron al ARN desnaturalizado. La reacción de la transcripción inversa se efectuó a 42°C durante 45 min, después se detuvo mediante incubación a 94°C durante 5 min. Después, los ADNc se conservaron a -80°C.

4- Amplificación génica

La amplificación de los ADNc se efectúa, de manera exponencial, por PCR (Reacción en Cadena de Polimerasa) (*Polymerase Chain Reaction*). Se utilizaron dos tipos de polimerasas: la polimerasa AmpliTaq Gold (*Thermophilus aquaticus*) (PE Applied Biosystems) y la polimerasa Pwo (*Pyrococcus woesei*) o Expand™ High Fidelity PCR System (Roche).

La amplificación se efectuó a partir de 5 µl de ADNc que se añadieron a una mezcla de reacción de PCR que contenía: 0,25 pmoles/µl de cebador sentido y antisentido (Tabla III), 200 µmoles de cada dNTP, 1,5 mM de MgCl₂, 1 U de AmpliTaq Gold ó 2,6 U de polimerasa Expand™ en presencia de los tampones PCR correspondientes. La reacción de PCR se realizó en un termociclador (PCR Sprint, Hybaid, Coger), de acuerdo con el siguiente protocolo: desnaturalización de híbridos ADNc – ARN a 94°C durante 9 min, seguido de 40 ciclos sucesivos, comprendiendo cada uno una desnaturalización de los ADN a 94°C durante 45 s, una hibridación de cebadores (900S/1280As o 6S/6As) a 58°C durante 30 s, síntesis de la cadena complementaria gracias a la polimerasa por elongación a 72°C durante 45 s. Finalmente, una elongación final a 72°C durante 4 min 30 s a 72°C.

Tabla III: Secuencias de cebadores utilizados para las reacciones de PCR y su posición en el genoma del HDV.

Cebadores	Secuencia 5' -> 3'	número de identificación	posiciones*
6S	gaggaaagaaggacgcgagacgcaa	SEC ID N°: 31	904-929
6As	acccctcgaaggtggatcga	SEC ID N°: 32	1141-1121
900S	catgccgacccgaagaggaaag	SEC ID N°: 33	889-911
1280As	gaaggaaggccctcgagaacaaga	SEC ID N°: 34	1289-1265
318S	ctccagaggacccctcagcgaac	SEC ID N°: 35	305-328
1150As	cccgcgggttggggatgtgaaccc	SEC ID N°: 36	1161-1138
960S	gtacactcgaggagtgaaggcg	SEC ID N°: 37	962-984
345As	tctgttcgctgaaggggtcct	SEC ID N°: 38	331-311

Cebadores	Secuencia 5' -> 3'	número de identificación	posiciones*
320S	ccagaggacccttcagcgaac	SEC ID N°: 39	307-328
420As	aacaccctcctgctagcccc	SEC ID N°: 40	446-427
R910S	ccggagttcctctcctcctcc	SEC ID N°: 41	1206-1227
R910AS	gttcgcgtcegagtccttcttte	SEC ID N°: 42	929-907
S1910R	gagctttcttcgattcggac	SEC ID N°: 43	1531-1550
AS1910R	gactgggccctcatgttcc	SEC ID N°: 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 viral HDV

La pareja de cebadores 6S y 6As permite amplificar un fragmento de ADN que corresponde al extremo carboxi-terminal del gen que codifica el antígeno delta.

La región R0 comprende el extremo carboxi-terminal del gen que codifica el Ag-HD y una parte de la región no codificante se ha ampliado para todas las muestras usando los cebadores 900S (SEC ID N°: 33) y 1280As (SEC ID N°: 34). El cebador 900S utilizado se ha deletado de 7 nucleótidos en el extremo 5', con respecto al utilizado por Casey et al., 1993a, citado anteriormente, para la clasificación de los genotipos del HDV.

La selección de estos cebadores permite, de manera sorprendente, la amplificación de un fragmento adecuado para permitir diferenciar los genotipos conocidos (I, II y III) de los nuevos genotipos.

Se obtuvieron las secuencias completas del genoma viral HDV de cuatro muestras (*dFr45*, *dFr47*, *dFr48* y *dFr73*) amplificando dos regiones solapantes R'1 (850 bases) y R'2 (1050 bases), respectivamente, usando los pares de cebadores 318S/1150AS y 960S/345AS.

Para la muestra *dFr644*, la variabilidad observada en la región correspondiente a los cebadores anteriormente descritos ha conducido a amplificar la región 644 (R644) usando un par de cebadores específico: 900S y 480AS.

Para la muestra *dFr910*, la secuencia de nucleótidos R0, ha permitido definir cebadores nuevos y específicos de la muestra para amplificar el genoma completo. Se seleccionaron dos pares de cebadores: los cebadores R910S y R910AS, que amplifican un fragmento de 1400 bases que corresponden a la región G910. Para cubrir la totalidad del genoma, ha sido indispensable otro par de cebadores S 1910R y AS 1910 R que amplifican un fragmento de 650 bases (región p910).

La amplificación de las diferentes regiones (R1, R2, R3, R644, R'1, R'2, G910 y p910) se efectuó como se ha descrito anteriormente. En la Tabla IV, se indican las temperaturas de hibridación y de elongación, así como el tiempo de elongación utilizados para cada una de las PCR.

Tabla IV: Amplificación de diferentes fragmentos del genoma

Regiones amplificadas	Tamaño de los fragmentos (bases)	Temperatura de hibridación (°C)	Temperatura de elongación (°C)	Temperatura 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

Se depositó un volumen de 8 µl de producto PCR en presencia de 2 µl de solución de depósito en un gel de agarosa al 1,3% preparado en un tampón Tris-Borato/EDTA 0,5X y que contenía 0,5 µg/ml de bromuro de etidio (BET). La electroforesis se efectuó en un tampón TBE 0,5X. La migración se efectuó, en presencia de un marcador de tamaño (Raoul™, Appligène). El fragmento amplificado se visualizó con rayos ultravioleta a 312 nm y se fotografió.

6- Clonación y secuenciación de los genomas del HDV

Antes de la etapa de clonación y secuenciación, los productos de amplificación se purificaron para eliminar cualquier traza de sales y de enzimas.

6.1.- Elongación mediante Taq-polimerasa convencional

Esta etapa se efectuó en el caso de haber realizado la amplificación del producto con la polimerasa Pwo. Esta permite añadir los restos de desoxiadenosinas (A) en los extremos 3' de los productos de la PCR, ya que la polimerasa Pwo, dotada de actividad exonucleásica 5' → 3', disminuye la incorporación de desoxiadenosinas.

Se añadió un volumen de 10 µl de ADN purificado a una mezcla de reacción de 70 µl que contenía: 0,2 mM de dNTP, 1,5 mM de MgCl₂, tampón 1X y 2,5 U de Taq Polimerasa (Perkin Elmer). La elongación se efectuó a 72°C, durante 30 minutos. Los productos de PCR se sometieron entonces a una nueva purificación con fenol-cloroformo y a una precipitación con etanol y después se recogieron en 10 µl de agua estéril.

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

La clonación se utilizó para confirmar la secuencia de nucleótidos del ADN amplificado. Se realizó usando el vector pCRII (Invitrogen).

El vector pCRII se presenta en forma lineal. Posee restos de desoxitimidina (T) que permite la clonación del producto amplificado gracias a los restos complementarios de desoxiadenosina (A) añadidos por la Taq polimerasa. También posee las secuencias promotoras Sp6 y T7, dos sitios de restricción *EcoRI* que flanquean el sitio de inserción del producto de la PCR y genes de resistencia a

ampicilina y a kanamicina. Una fracción del gen *lacZα*, que codifica la β-galactosidasa, facilita la selección de recombinantes por el color de las colonias. De hecho, los plásmidos que tienen integrado el inserto no expresan el gen *lacZα*. Las colonias bacterianas son por tanto blancas en presencia del sustrato β-galactosidasa (X-Gal o 5-bromo 4-cloro 3-indoil β-galactósido, Roche) y de un inductor del gen (IPTG o isopropil-tio β-D-galactósido, Roche). De esta manera, las bacterias recombinantes se seleccionan gracias a su resistencia a la ampicilina y a una selección azul-blanca.

La proporción molecular inserto/vector seleccionada es de 3/1 y el volumen del producto PCR utilizado es variable, según la cantidad de ADN calculado por electroforesis en gel de agarosa como se ha descrito anteriormente. La mezcla de reacción de 10 µl contiene 50 ng de vector pCRII, la cantidad de inserto correspondiente, 4 U de T4 ADN Ligasa, y el tampón Ligasa 1X. La reacción de ligamiento se realiza durante 18 horas a 14°C. Después los tubos se conservan a +4°C.

Las bacterias *Escherichia coli* TOP10F' (Invitrogen), que se han hecho competentes por tratamiento con cloruro de calcio, se conservan a -80°C, preparadas para el uso. Un volumen de 50 µl de bacterias competentes se pone en contacto con 3 µl de la solución de ligamiento durante 30 minutos, en hielo. Un choque térmico (30 s a 42°C) hace que el ADN plasmídico penetre en las bacterias, que se colocan de nuevo inmediatamente en hielo durante algunos minutos, antes de incubarse 1 hora a 37°C, en 250 µl de medio SOC (Tryptona al 2%, NaCl 10 mM, KCl 2,5 mM; MgCl₂ 10 mM; glucosa 20 mM, extractos de levadura 5 g/l). A continuación se aíslan las colonias en placas de Petri que contienen LB agar (medio de Luria-Bertani), complementado con ampicilina (50 µl/ml), 40 µl de X-Gal (40 mg/ml) y 40 µl de IPTG (100 mM) y se distribuyen.

6.3.- Extracción plasmídica y análisis del inserto

Las colonias blancas se sembraron en caldo LB-ampicilina (50 µl/ml) y se incubaron durante 18 horas, a 37°C, con agitación. Se conservó una colonia azul, es decir sin fragmento insertado, como control negativo de ligamiento.

La extracción plasmídica se realizó usando un kit comercial QIAprep® Spin Miniprep (Qiagen). En breve, después de la centrifugación (3000 revoluciones por minuto a +4°C) y la eliminación del sobrenadante, el sedimento bacteriano se suspendió en 250 µl de tampón (Tris-HCl 50 mM pH 8, EDTA 10 mM, RNasa A 100 µl/ml) y se realizó la lisis añadiendo 250 µl de tampón alcalino ((NaOH 200 mM, SDS al 1 %). Después de 5 min de homogeneización, se añadieron 350 µl de tampón (acetato de potasio 3 M, pH 5,5). A continuación, el sobrenadante que contenía el ADN plasmídico se transfirió a una colonia QIAprep. Una centrifugación eliminó el eluado en el tubo colector.

La colonia se lavó usando un tampón de etanol, se secó, y después el ADN se eluyó en 50 µl de agua estéril.

Para verificar la inserción del fragmento de interés, el plásmido se digirió entonces mediante la enzima de restricción EcoRI. La digestión se efectuó en una mezcla de reacción de 30 µl que contenía: 2 µl de la solución plasmídica, 10 U de enzima EcoRI (Appligène) y tampón de reacción 1X. La digestión duró 2 horas a 37°C y el resultado se visualizó por electroforesis en gel de agarosa.

6.4.- Secuenciación por el método BigDye Terminator

La secuenciación se realizó sobre los productos PCR previamente purificados sobre colonias Microcon 50 (Amicon) o sobre el ADN plasmídico. Los fragmentos se secuenciaron directamente con los cebadores de PCR (fragmento R0 secuenciado con los cebadores 900S y 1280As) o después de la clonación en el vector pCRII utilizando cebadores universales (Sp6 y T7).

Para cada una de las regiones amplificadas se conservaron dos clones diferentes, con objeto de eliminar eventuales ambigüedades durante la lectura de las secuencias nucleotídicas.

La secuenciación se efectuó usando el reactivo BigDye Terminator (PE, Applied Biosystems). El principio de secuenciación consiste en una electroforesis vertical en gel de poliacrilamida del ADN marcado por cuatro fluorocromos diferentes. Las matrices del ADN se depositan sobre el gel y se separan según su tamaño, antes de someter el gel a un haz de láser continuo. El láser excita los fluorocromos emitiendo cada uno a una longitud de onda diferente, detectada por un espectrógrafo. Un programa informático, acoplado al secuenciador, permite el análisis automático y la conversión de los datos en secuencias nucleotídicas.

La mezcla de reacción de 10 µl comprende: 4 µl de la solución de marcado (desoxinucleótidos trifosfato (dATP, dCTP, dGTP, dUTP), AmpliTaq ADN polimerasa, MgCl₂, Tampón Tris-HCl pH 9), 20 pmoles de cebador (sentido o antisentido) y 500 ng de plásmido purificado en las colonias de tipo Centricon. Las reacciones de las secuencias (sentido y antisentido) se efectúan 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). Los productos se precipitan a continuación en 40 µl de etanol al 70%, se depositan en gel y se analizan usando un secuenciador automático de tipo ABI PRISM 377.

Las secuencias brutas obtenidas se presentan en forma de electroforegramas. La validación y el aprovechamiento de las secuencias se realizan usando el programa informático Séquence Navigator (PE, Applied Biosystems). Con la finalidad de minimizar errores, éstas son objeto al menos de una doble lectura, por dos cebadores de secuenciación diferentes (Sentido y Antisentido).

A continuación estas secuencias se introducen en un ordenador utilizando el programa informático *DNA Strider 1.3* que permite un análisis rápido de las secuencias.

7.- Análisis informático de las secuencias nucleotídicas y proteicas

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

Las secuencias obtenidas (22 secuencias) se han comparado con 21 secuencias genómicas completas del HDV disponibles en el GenBank (Tabla V)

Tabla V: Números de acceso de diferentes aislados

Números de acceso (GenBank)	Nombre del aislado	Origen geográfico
1 X04451	Italia 1 (A20)	Italia
2 M84917	Lebanon 1	Líbano
3 X85253	Paciente A.	Cagliari (Italia)

Números de acceso (GenBank)	Nombre del aislado	Origen geográfico
4 X60193	18/7/83 (paciente S)	(paciente S) Japón
5 M92448	Taiwán	Taiwán
6 L22061	Columbia	Colombia
7 X77627	Suero humano	China central
	Chino	
8 L22064	Perú-2	Perú
9 L22063	Perú-1	Perú
10 L22066	US-2	Estados Unidos
11 M58629	Nauru	Isla de Nauru
12 U81988	Somalia	Somalia
13 U81989	Etiopía 1	Etiopía
14 AF098261	Canadá	Canadá (Quebec)
15 U19598	Taiwán 3	Taiwán
16 AF018077	TW2b	Taiwán
17 L22062	Japón 3	Japón
18 AF309420	Miyako	Isla de Miyako (Okinawa, Japón)
19 D01075	Us- 1	Estados Unidos
20 M21012	W15	Transmisión experimental (Marmotte)
21 AJ307077	W5	Transmisión experimental (Marmotte)
22 AJ309868 a AJ309881	Aislados de Yakouti	Yakouti (Rusia)

La primera etapa consiste, en general, en alinear las secuencias de interés con las secuencias HDV de referencia descritas y catalogadas en el banco de datos (Genbank), utilizando el programa CLUSTAL W1.8 (Thompson et al., N.A.R, 1994, 22, 4673-4680). Han sido necesarias correcciones manuales menores usando el programa informático *SeqPup* para optimizar el alineamiento.

Se han seguido dos enfoques: la utilización del alineamiento de proteínas para el gen HD y el estudio de la estabilidad de las posiciones alineadas usando un programa informático de alineamiento adaptado.

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

- enfoque fenético (distancia genética)

El principio de este método es encontrar pares de secuencias vecinas, minimizando la longitud total de las ramas del árbol. Este enfoque permite reconstruir una filogenia sobre la base de cálculo de la similitud global entre las secuencias comparadas dos a dos que se expresa por una distancia. Es un método que permite convertir los datos de las secuencias en valores numéricos de distancias, dispuestos en matrices. La topología del árbol se construye para reagrupar las secuencias que poseen la mayoría de los caracteres en común, utilizando uno de los métodos de reagrupamiento como el método del vecino más próximo o “*neighbor-joining*” (Saitou et al., 1987).

- enfoque cladístico (máxima parsimonia)

El principio de este método consiste en establecer relaciones de parentesco entre las secuencias mediante la búsqueda de bases nucleotídicas compartidas, minimizando los sucesos genéticos. El algoritmo de máxima parsimonia construye un árbol filogenético de tal manera que implica el mínimo de mutaciones. El árbol aceptado es el que necesita menos cambios. Este método es sensible a las diferentes tasas de mutaciones a lo largo de las ramas. Los grupos de secuencias que comparten un ancestro común constituyen los “clados” o “grupos monofiléticos, excluyendo cualquier otra secuencia.

- enfoque estadístico

El método de máxima verosimilitud se considera como un enfoque estadístico. El programa calcula la probabilidad para que una secuencia evolucione hacia otra a lo largo del tiempo. En otras palabras, se trata de considerar los cambios a nivel de cada sitio o carácter como sucesos probabilísticos independientes. Este algoritmo de verosimilitud se acumula en todos los sitios y la suma se maximiza para calcular la longitud de la rama del árbol. Este método necesita un tiempo de cálculo considerable para buscar el árbol filogenético, el más verosímil, que corresponde a las secuencias observadas de manera que tiene en cuenta la probabilidad de cambio de cada carácter.

Todos los análisis filogenéticos se han realizado utilizando los programas informáticos *Phylip 3.75 (PHY Logenetic Inference Package)* (Felsenstein et al., 1989) y *Paup * versión 4. Obêta6* (Phylogenetic Analysis Using Parsimony) (Swofford et al., 1998).

El análisis de distancia se ha calculado por el método de Kimura de dos parámetros, que considera como diferentes los índices de transición (mutaciones T <-> C et G <-> A) en cada sitio, y de transversión (mutaciones « A o G » <---> « T o C ») en cada sitio.

La fiabilidad y la robustez de los grupos de secuencias (o topologías) se evalúan estadísticamente por el enfoque de remuestreo (o bootstrap) sobre 10³ y 10⁴ remuestreos.

Los resultados obtenidos se presentan 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). También permite visualizar los valores de remuestreo (bootstrap), en cada nudo y determinar un taxón como grupo externo (secuencias del genotipo-III).

La traducción en aminoácidos del gen delta se efectúa usando el programa *DNAStrider* versión 1.3. El alineamiento de las secuencias proteicas se realiza como se ha descrito anteriormente.

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

El genotipado del HDV se efectúa por PCR-RFLP de la región R0, según las siguientes etapas:

- Etapas 1: Las dos enzimas de restricción *SmaI* y *XhoI* (New England Biolabs) digieren los productos PCR: 5 U de enzimas *SmaI* o *XhoI*, digieren por separado, en dos tubos, 10 µl de producto amplificado, a 30°C y 37°C, respectivamente durante 3 horas en un volumen final de 50 µl en presencia de tampón apropiado y agua estéril. Los productos de digestión se visualizan con rayos ultravioleta como se ha descrito anteriormente y se determinan los tamaños de los fragmentos por comparación con un marcador de tamaño (50 pb escala de ADN o marcadores V y VI, Life Technologies GibcoBRL).
- Etapas 2: Otra enzima, *SacII* (New England Biolabs), digiere durante 3 horas a 37°C, las muestras que presentan un perfil distinto al perfil del genotipo I, y los productos de digestión se visualizan como en la etapa 1.
- Etapas 3: el genotipo del virus se determina a partir del análisis de la combinación de perfiles de restricción *SmaI*, *XhoI* y *SacI*.

9- Algoritmo de genotipado del HDV por PCR-RFLP

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

- la primera, consiste en la escisión del fragmento R0, amplificado por RT-PCR a partir de los ARN extraídos de sueros de pacientes, por dos enzimas de restricción, *SmaI* y *XhoI*;
- la segunda, para los pacientes de “perfil no-I”, consiste en una escisión de R0 por la enzima *SacII*;
- la secuenciación de la región R0 o de la región que codifica la p24 (o si es necesario la totalidad del genoma), seguido de análisis filogenético sólo se realizará en segundo lugar si se obtienen perfiles de restricción no habituales.

EJEMPLO 2: DEMOSTRACIÓN DE NUEVOS GENOTIPOS DEL HDV

1- Análisis filogenético de la región R0

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

El estudio filogenético se efectuó a partir de un alineamiento de secuencias de 336 bases de R0 (las regiones ambiguas se eliminaron), incluyendo, además de las 22 secuencias estudiadas, 15 secuencias de referencia y 6 secuencias R0 del HDV de Yakouti (Pt13, 26 (SEC ID N°: 66), 29, 62 (SEC ID N°: 67), 63 y 704). El nombre proporcionado a las secuencias corresponde a dFr (por “delta Francia”) seguido de un número de suero del paciente.

a) análisis de la distancia genética

En la figura 1 se presenta el árbol filogenético, obtenido por reconstrucción a partir de distancias genéticas de la región R0.

La topología del árbol individualiza los genotipos I y III, representados respectivamente por siete y tres secuencias nucleotídicas de referencia. Las otras secuencias de referencia se representan

por las secuencias de tipo II (Japón, Taiwán-3 y las secuencias de Yakouti) y un grupo de dos secuencias (TW2b, Miyako) descritas respectivamente cada una como prototipo de « sub-tipos IIB y IIC ».

Este árbol demuestra que las secuencias virales que provienen de 22 muestras analizadas corresponden a dos situaciones:

- 5 - 11 secuencias están afiliadas a las secuencias del genotipo I, a 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 manera heterogénea dentro del genotipo-L.
- las 11 secuencias restantes están muy alejadas del genotipo-I y del genotipo-III. Además, ninguna de estas secuencias se reagrupa directamente con las secuencias de tipo-II (Japón, Taiwán-3, Yakouti 13, 26, 29, 62, 63, 704) o con el grupo de secuencias (TW2b, Miyako); estas secuencias de
- 10 referencia forman en sí mismas dos grupos distintos.

La topología del árbol, obtenido por reconstrucción a partir de distancias genéticas de la región R0, muestra que las moléculas nucleicas aisladas a partir de diversos HDV variantes se distribuyen dentro de cuatro sub-grupos (figura 1):

- 15 la molécula dFr644 que aparece aislada; posee sin embargo, con un grupo de tres moléculas (dFr45, dFr2066 y dFr1843), un nudo que se sostiene por un valor de remuestreo de solamente el 66,7%,
- en cambio, la rama que une las moléculas dFr45, dFr2066, dFr1843 es robusta, al sostenerse por un valor de remuestreo (BV) del 99,9%.
- un conjunto de cinco moléculas: dFr47, dFr910, dFr69, dFr73 y dFr1953 se sostiene por un valor BV
- 20 del 100% y
- un par de moléculas dFr48 y dFr2020, que también se sostiene por un BV del 100 %.

b) análisis de máxima parsimonia

En la figura 2 se presenta el árbol filogenético, obtenido por reconstrucción a partir de la máxima parsimonia de la región R0.

- 25 El análisis de máxima de parsimonia sostiene la misma topología que el análisis de la distancia genética. La reconstrucción demuestra la existencia dentro de 11 secuencias variantes, de los mismos tres grupos monofiléticos; por ejemplo, con este enfoque, el grupo de cinco moléculas *dFr47*, *dFr910*, *dFr69*, *dFr73* y *dFr1953* también se sostiene por un BV del 97% (Figura 2).

- Las 11 moléculas variantes, las moléculas del genotipo II y el conjunto [TW2b, Miyako]
- 30 parecen derivar de una rama común que podría, por comparación con las secuencias del genotipo I y del genotipo III, individualizar el conjunto de las secuencias del genotipo II. Por lo tanto, los valores de remuestreo que sostienen esta ramificación son relativamente modestos: 88,5 % en NJ y 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 dentro de
- 35 las 11HDV variantes o entre estas variantes y las secuencias del genotipo II, parece más elevado que entre el conjunto de aislados del genotipo I o dentro de las tres moléculas que definen el genotipo III.

El conjunto de estos resultados destacan la caracterización de nuevos genotipos del HDV.

2 – **Análisis filogenético de la totalidad de los genomas**

a) Reconstrucción del genoma completo a partir de fragmentos amplificados

Con objeto de estudiar el genoma completo de estas variantes y, en el caso de precisar su afiliación, se amplificaron varias regiones del genoma HDV (Tabla II) a partir de 6 muestras que incluyen al menos un miembro de cada uno de los 4 sub-grupos y se seleccionaron tres representantes del grupo mayoritario: *dFr45*, *dFr47*, *dFr48*, *dFr73*, *dFr644* y *dFr910*.

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

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

Todas estas regiones genómicas amplificadas se han clonado en un vector pCRII (Tabla IV). Se han secuenciado dos clones correspondientes a cada uno de los fragmentos amplificados. La reconstrucción de las secuencias completas consenso de ADNc del HDV se ha efectuado después del alineamiento de las regiones solapantes 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 4	-	-	R910 clon R910 clon 31	29 910R1 clon 4 910R1 clon 5

b) Análisis de seis nuevas secuencias genómicas completas del HDV, de origen africanob₁) Características clínicas de 6 pacientes

Cinco pacientes son originarios del África del Oeste, y un paciente ha pasado temporadas en Camerún. En el momento de la extracción, esos pacientes residían en la región parisina desde al menos dos años. Todos estos pacientes padecían hepatitis grave y en la figura 3 se resumen los datos clínicos.

b₂) Organización genómica de nuevas secuencias del HDV

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

El tamaño de los genomas completos es diferente en las seis secuencias de los seis aislados del HDV de origen africano, lo que confirma la variabilidad del HDV:

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

El análisis después del alineamiento de las diferentes secuencias estudiadas, revela un grado elevado de conservación en las regiones del genoma HDV correspondiente a las ribozimas responsables de las escisiones de los ARN genómicos y antígenómicos. Asimismo, la fase de lectura que codifica el antígeno delta se encuentra en la cadena antígenómica. Un codón triptófano (UGG) es el único caracterizado por dos secuencias (*dFr47*, *dFr910*) y una ambigüedad (G/A) encontrada para las otras cuatro secuencias indica que probablemente la pequeña y gran proteína delta están muy probablemente sintetizadas. Las regiones variables comprenden la parte no codificante así como los extremos 5' y 3' del gen *LDH*. De manera notable, 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 pseudo doble cadena (en la posición 797 de la secuencia de referencia italiana (Wang et al., 1987)).

c) Comparación de seis secuencias HDV de origen africano con las secuencias representativas de diferentes genotipos

La comparación de seis nuevas moléculas con las moléculas conocidas, representativas de tres genotipos conocidos, indica una similitud nucleotídica que se sitúa entre el 71,7% (*dFr45* frente a Líbano) y el 80,0% (*dFr73* frente a Yakouti p26) en relación con las moléculas de los genotipos I, II y las moléculas TW2b y Miyako. De hecho, para cada uno de los seis aislados, el promedio de similitud nucleotídica es del orden del 73,3% al 74,6% con las moléculas del genotipo I, del 74,5% al 78,8% con las del genotipo II y del orden del 74,6% al 77,8% con las moléculas TW2b/Miyako. En cambio, la similitud nucleotídica con el aislado de Perú (genotipo III), sólo es del 63,9% al 66,0% lo que confirma el distanciamiento particular de esta molécula (Tabla VII). Además, cuando se comparan entre sí, las seis moléculas que corresponden a estos genomas completos y que definen las seis variantes *dFr45*, *dFr47*, *dFr48*, *dFr73*, *dFr644* y *dFr910*, sólo el grupo de moléculas *dFr73*, *dFr910* y *dFr47* presentan una similitud de secuencia del orden del 90%. Las moléculas de *dFr45*, *dFr48* y *dFr644* también están distanciadas entre sí con respecto a los genotipos I, II, las secuencias TW2b / Miyako y el grupo de moléculas *dFr73*, *dFr910* y *dFr47* (del orden del 73,2% al 78%) (Tabla VIII).

Tabla VII: Porcentaje de similitud de las secuencias completas HDV africanas con los diferentes genotipos conocidos (Cálculo de la media).

Aislados HDV*	Tipo I	Tipo II	TW2b/Miyako	Tipo III
<i>dFr45</i>	73,3 71,7 - 74,6	74,5 73,2 - 75,5	74,6	66
<i>dFr47</i>	74,2 73,0 - 75,0	78,6 78,2 - 79,9	77,4	65,5
<i>dFr48</i>	73,3 72,0 - 74,0	77,1 76,6 - 77,7	75,5 74,4 - 76,6	65,4
<i>dFr73</i>	74,1 73,0 - 75,0	78,8 77,7 - 80,0	77,8 77,5 - 78,0	65,9
<i>dFr644</i>	73,6 72,2 - 74,6	76,8 76,2 - 77,2	77,0 76,9 - 77,2	63,9
<i>dFr910</i>	74,6 73,0 - 75,8	77,9 77,0 - 78,6	77,2 77,0 - 77,5	64,6
* Los aislados HDV de referencia corresponden a los genomas completos estudiados en el ejemplo 2.1.				

Tabla VIII: Porcentaje de similitud de las nuevas moléculas del HDV entre sí.

	<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 seis moléculas del HDV de origen africano y moléculas representativas de diferentes genotipos

El análisis filogenético se ha realizado en las seis secuencias completas de origen africano, dieciséis secuencias de referencia y dos secuencias de Yakouti (Pt26 y Pt62). La Figura 4 ilustra los resultados obtenidos por análisis de distancia. El árbol filogenético reconstruido por el "vecino más próximo" ("Neighbor-Joining") (NJ) muestra que ninguna de las seis secuencias estudiadas (*dFr45*, *dFr47*, *dFr48*, *dFr73*, *dFr644* y *dFr910*) se reagrupa con las secuencias de referencia de genotipo I o de genotipo III. La afiliación de estas secuencias africanas con las secuencias de los genotipos II (de las secuencias TW2b y Miyako descritas respectivamente como los sub-tipos IIB y IIC) no se sostiene, por valores de remuestreo elevados (<70%) (Wu et al., 1998). Además, las secuencias TW2b y Miyako parecen formar un grupo distinto y monofilético con un BV del 100%. Estas dos secuencias parecen constituir por sí mismas un "clado" que representa un genotipo diferente del tipo II.

En los análisis de distancia, las seis secuencias africanas se subdividen en 3 subgrupos distintos (soportados por BV superiores al 90,3% para 10⁴ remuestreos). Las secuencias *dFr47*, *dFr73* y *dFr910* constituyen un grupo cuya rama se basa en un valor de bootstrap del 100%. Para reafirmar estos resultados, se realizó el estudio de máxima parsimonia en el mismo juego de secuencias (Figura 5).

Arraigando, el árbol de manera artificial gracias a la secuencia 'Perú 1', se individualiza, como en todos los análisis realizados anteriormente, el conjunto de secuencias del genotipo-I (BV = 100%). La topología de otras secuencias sostiene la distribución de los aislados africanos y asiáticos en varios grupos; lo que muestra el interés de la utilización de la región R0. El genotipo II reagrupa las secuencias de Yakouti, Taiwán-3 y Japón con un BV del 99,9 % en 10⁴ remuestreos. Asimismo, 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*, *dFr910* (BV = 100 %) sostiene la afiliación de estas secuencias en un subgrupo. En cambio, la secuencia *dFr48*, que posee con los aislados del grupo anterior (*dFr910*, *dFr47*, *dFr73*) una similitud nucleotídica respectiva de 76,1, 77,1 y 77,7% se reagrupa con estas secuencias solamente en el 55,4% de los remuestreos, lo que sugiere su posible individualización. Aunque que aparecen distantes entre sí, el grupo *dFr45* y *dFr644*, se observa con un BV elevado (NJ = 96,5 / MP = 88,6) en el contexto estudiado.

Por consiguiente, a la vez que los análisis filogenéticos de las regiones R0 y las secuencias completas de las secuencias africanas indican que los grupos difieren entre sí y podrían constituir tres (incluso cuatro) genotipos distintos; estos resultados demuestran la existencia de al menos siete genotipos del HDV.

3- Análisis de la secuencia de aminoácidos (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 pequeña proteína delta (sHD) o forma p24. La gran proteína delta (LHD) o forma p27 posee el mismo extremo amino-terminal y una extensión de 19 a 20 aminoácidos en su extremo carboxi - terminal.

En la figura 6 se presenta el alineamiento de la secuencia del antígeno HD de seis secuencias africanas con las secuencias de las proteínas HD conocidas.

El análisis de secuencias demuestra que los seis aislados de origen africano tienen una identidad de aminoácidos del orden del 69 al 77% con las secuencias del genotipo I, del 71 al 79% con los aislados del genotipo II, del 72 al 78% con las secuencias TW2b / Miyako y del 63% con el aislado de Perú (genotipo III).

El tamaño de las proteínas que corresponden a los nuevos aislados varía entre 213 y 214 aminoácidos. Todas estas proteínas tienen el mismo perfil de hidrofobicidad. La forma p24 posee 2 pequeñas regiones hidrófobas, una situada cerca de los aminoácidos 50-60 (entre el sitio de polimerización y el SLN) y la otra entre las posiciones 160 y 172 (enfrente de un motivo extremadamente conservado). Entre los diferentes genotipos otros dos dominios se encuentran bien conservados: se trata del dominio de fijación al ARN y del dominio de localización nuclear. Como se ha descrito en la bibliografía, en el extremo carboxi-terminal de la proteína delta (entre los aminoácidos 195 y 215) constituye una región hipervariable. Solamente se conservan dos aminoácidos de los 19-20. Se trata de la cisteína (C) que corresponde al sitio de farnesilación de la forma grande de la proteína HD y la glicina (G) carboxi-terminal. Además, se encuentran las secuencias firma específicas de los aislados del

mismo genotipo, por ejemplo, los 19 aminoácidos específicos de la proteína grande del genotipo I o los 20 aminoácidos del genotipo III.

En cambio, para las secuencias proteicas de los aislados de origen africano, los aislados del genotipo II y de TW2b/Miyako, el extremo carboxi-terminal parece subdividido en dos dominios. El dominio variable se representa por los aminoácidos 197 a 205 y el dominio conservado que va de los aminoácidos 206 a 215 (RLPLLECTPQ) (Figura 5).

4- Definición de 5 clados del HDV

El análisis de secuencias completas de seis aislados de África permite definir siete clados del HDV correspondientes a los siguientes genotipos (Tabla IX):

Tabla IX: Correspondencia Clado/Genotipo

Clado	Genotipo	Aislados
1	I	Italia, W5, W15, Us 1, Us2, Líbano, Etiopía, Somalia, Nauru, China, Cagliari, Canadá...
2	IIA	Japón, Taiwán ³ , Yakouti 26, Yakouti 62
3	III	Perú 1
4	IIB, IIC	TW2b, Miyako
5	?	dFr91 0, dFr73, dFr47
6	?	dFr 48
7	?	dFr45, dFr644

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

El genotipado se efectúa de acuerdo con el protocolo descrito en el ejemplo 1.8.

1- Pérdida de sensibilidad de la PCR 6A/6S

Inicialmente, tres pacientes Ag-HBs positivos, han planteado un problema de diagnóstico de infección delta. De hecho en estos pacientes se observa una hepatitis grave asociada con la presencia de IgM anti-HDV pero una ausencia de replicación del HDV por RT-PCR usando los cebadores 6A-6S descritos en Deny et al. (1991, 1993, 1994, citados anteriormente) para el diagnóstico rutinario de la infección por HDV. La PCR 6A/6S amplifica un fragmento del ADNc de 234 pb que corresponden a la parte del extremo carboxi-terminal del gen LHD (posiciones 904 a 1141 en el genoma viral).

Los ARN extraídos del suero de estos mismos pacientes se reamplificaron usando un par de cebadores 900S y 1280AS que definen la región R0.

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

Estos resultados se confirmaron en una serie de muestras de suero de pacientes que se analizaron en paralelo con el par de cebadores 6A-6S y 900S-1280AS. En 286 muestras, 14 son positivas únicamente con la PCR de R0.

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

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

Los métodos clásicamente utilizados de PCR-*RFLP* (Wu et al., 1995a; Wu et al., 1995b; Casey et al., 1996) permiten distinguir tres genotipos delta diferentes. La utilización de la enzima de restricción *SmaI* no diferencia todos los genotipos -I, -IIA, -IIB reconocidos actualmente y la enzima *XhoI* se ha utilizado para diferenciar el «sub-tipo IIA» del «sub-tipo IIB» (Wu et al., 1995b).

La asociación de dos enzimas *SmaI* y *XhoI* en una primera etapa revela siete perfiles distintos (de P1 a P7) (Tabla X). Estos siete perfiles no se superponen exactamente con los siete clados (HDV-1 a HDV-7). Como consecuencia, las muestras de perfil «no-P1» se escinden en una segunda etapa mediante la enzima *SacII*, dando como resultado de esta manera la obtención, de manera combinada, de diez perfiles delta distintos (de D1 a D10) (Tabla XI) que pueden unirse específicamente con los diferentes clados descritos, gracias a los análisis de filogenia.

Tabla X: Perfiles de restricción escindidos de la Región R0, por las enzimas *SmaI* et *XhoI*.

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

Tabla XI: Perfiles de restricción esperados después de la escisión de la Región R0, por las enzimas *SmaI*, *XhoI* y *SacII*.

ETAPA	2	Genotipos	Fragmentos	Perfil	Perfil	Combinado
descritos			<i>SacII</i>	<i>SacII</i>	<i>SmaI-XhoI/SacII</i>	
			Tamaño (pb)			
I			362,38	<i>sc1</i>	<i>S1 X1 Sc1</i>	D1
IIA			266,92,38	<i>Sc2</i>	<i>S2 X2 Sc2</i>	D2
IIB			268,130	<i>Sc3</i>	<i>S2 X3 Sc3</i>	D3
IIC (Miyako)			268,130	<i>Sc3</i>	<i>S2 X4 Sc3</i>	D4

ETAPA descritos	2	Genotipos	Fragmentos <i>SacII</i> Tamaño (pb)	Perfil <i>SacII</i>	Perfil <i>SmaI-XhoI/SacII</i>	Combinado
III			405	<i>Sc4</i>	<i>S3 X5 Sc4</i>	D5
II Yakouti			266,92,38	<i>Sc2</i>	<i>S4 X2 Sc2</i>	D6
<i>dFr45</i>			268,130	<i>Sc3</i>	<i>S1 X2 Sc3</i>	D7
<i>dFr644</i>			397	<i>Sc4</i>	<i>S1 X2 Sc4</i>	D8
<i>dFr47, 73, 910</i>			268,130	<i>Sc3</i>	<i>S4 X2 Sc3</i>	D9
<i>dFr48</i>			268,130	<i>Sc3</i>	<i>S2 X2 Sc3</i>	D10

3- Genotipado de muestras de pacientes por PCR-RFLP

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

- no se ha encontrado ningún genotipo-II ó -III.
- el 89,7% de los pacientes presentaban un perfil D1 (genotipo I) y el 10,3% un perfil « no-I »
- 5 - se detectaron dos nuevos perfiles *XhoI* (X6 y X7) que conllevan tres combinaciones nuevas (D11, D12 y D13) complementarias (Tablas XII y XIII).

Tabla XII: Nuevos perfiles de restricción *XhoI* obtenidos en cinco pacientes originarios de África del Oeste

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

Tabla XIII: Nuevos perfiles de restricción *XhoI*, *SmaI*, *SacII*, obtenidos en cinco pacientes originarios de África del Oeste

10

ETAPA PACIENTES	2	Fragmentos <i>SacII</i> Tamaño (pb)	Perfi I <i>SacII</i>	Perfil Combinado (<i>SmaI-XhoI</i> / <i>SacII</i>)
dFr1843		267, 130	<i>Sc3</i>	<i>S1 X2 Sc3</i> D7
dFr1953		267, 92, 38	<i>Sc2</i>	<i>S1 X2 Sc2</i> D11
dFr2020		262, 130	<i>Sc3</i>	<i>S2 X2 Sc3</i> D10
dFr2088		396	<i>Sc4</i>	<i>S1 X6 Sc4</i> D12
dFr2066		396	<i>Sc4</i>	<i>S1 X7 Sc4</i> D13

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

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

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

REFERENCIAS BIBLIOGRÁFICAS

- Casey J.L. et al., Proc. Natl. Acad. Sci. USA, 1993a, 90, 9016-20.
- 5 • 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.
- 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.
- 10 • Deny P. et al., J. Gen. Virol., 1991, 72, 735-9.
- Felsenstein J. et al., Cladistics, 1989, 5, 164-166.
- 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.
- 15 • Imazeki F. et al., J. Virol., 1990, 64, 5594-5599.
- Imazeki F. et al., Nucl. Acid Res., 1991, 19, 5439-5440.
- Lai M.M.C. et al., S. Hadziyannis, J. Taylor et F. Bonino (ed.), Hepatitis delta virus. Molecular biology, Pathogenesis, and Clinical aspects, 1993, 382, 21-27. Wiley-Liss, New York.
- Lee C.M. et al., Virology, 1992, 188, 265-273.
- 20 • 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.
- Roingear P. et al., Clin. Infect. Dis., 1992, 14, 510-14.
- Saitou N. et al., Mol. Biol. Evol., 1987, 4, 406-25.
- 5 • 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.
- Thompson J.D. et al., Nuc. Acid. Res., 1994, 22, 4673-80.
- 10 • 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.
- Wu J.C. et al., Lancet, 1995b, 346, 939-41.

LISTA DE SECUENCIAS

<110> ASISTENCIA PÚBLICA-HOSPITAL DE PARIS

DENY Paul

5 RADJEF Nadja

HUC-ANAIS Patricia

<120>.

10 <130> S1020FR17

<140>

<141>

15 <160> 67

<170> PatentIn Ver. 2.1

<210> 1

<211> 1672

20 <212> ADN

<213> virus de la hepatitis D

<220>

<221> cepa

25 <222> (1)..(1672)

<223> secuencia genómica completa de dFr45

<400> 1

```

atgagccaag tccgacgaag agtccggaga tgggagagga gagagacccc cgaggaggca 60
aagtcgagag ggcggagaca ctctgaggag aggactccca agaggacaag aaggactcaa 120
gatccgggaa cgaatcccca aaacgccggt aaagctcttg gaatgaggaa aacgggaggg 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 gcccccggtt ggctccgagt gaggcttatc 480
ccggggaatt tggctawctt ccccaactag ccggtcccg atccccctcc aratgaacgg 540
ggggactccg gcctctcggg gaatcggcgg ctccatggtg aacccccgca gccccccact 600
ccacactcct tccccccgcg gggcccccca tccaaagatg gaactccacc cctcagggtt 660
cgcyatccta cccttttctt acctttggcc ggcatggtcc cagcctctc gctggcgccg 720
gctgggcaac attccgaggg gaccgtccct cggcaatggc gaatgggacc caagacctcc 780
aagattccca gagagaatct ggaggtgact gggcactccc tttgccatcc gagtggacgt 840
tcgtcctcct tcggatgccc aggtcggacc acggggaggt ggagatgcca tgccgaccgg 900
aagaggaaag aaggattcgg acgcaaacct gtgagtggaa gttcgctctt tattgggggg 960
tacactcgag gagtggaaag cggggagggg gggtcggact aggtccctat ggaaactgtc 1020
ggtctcctcg gatgtcgagt cctctccccg ttctggagaa gggggactcc gggacaccta 1080
gcagttgagg aacgaagccg cccccggcg ctcccctcg tggccctcg ggagggttca 1140
catccccaag ccgcggggcg gctattcttc ttgctctttg ctcgtcacg gcggtcaact 1200
tcctgagttc ctctttcttc tccttgcgga gggactttcc tcctgcggac agctgcttct 1260
tcttgttctc gagggccttc cttcgtcggg gageccgtct ctctcgtcg gtgaagcctc 1320
ccttgttact ctttttccct ggtccggagt cgacctccat ccgatctgtt cgggggtctc 1380
tcgcccgggg agctccctcc ccagtccga cctttccaat tattcctttg atgtttccca 1440
gccagggatt gtgcgtctca agtttcttga tgggtcttct tgtcttcgg agcctcctct 1500
cgagttcttc cgcgtcttcc cttgctttgw cccacttcga aagggtctcc tcccttccct 1560
tccgggatct cttcgtatcg gcgtggctca tctcggcaga ggcgggcgat cctcagttct 1620
cttactcttt tcttgaaaga ggagactgct ggacgctccg ccgtgtccga gc 1672

```

<210> 2

<211> 642

5 <212> ADN

<213> virus de la hepatitis D

<220>

<221> CDS

10 <222> (1)..(692)

<223> ADNc de LHD

<400> 2

atg agc cac gcc gat acg aag aga tcc cgg aag gga agg gag gag acc	48
Met Ser His Ala Asp Thr Lys Arg Ser Arg Lys Gly Arg Glu Glu Thr	
1 5 10 15	
ctt tcg aag tgg gwc aaa gca agg gaa gac gcg gaa gaa ctc gag agg	96
Leu Ser Lys Trp Xaa Lys Ala Arg Glu Asp Ala Glu Glu Leu Glu Arg	
20 25 30	
agg ctc cgg aag aca aag aag acc atc aag aaa ctt gag gac gac aat	144
Arg Leu Arg Lys Thr Lys Lys Thr Ile Lys Lys Leu Glu Asp Asp Asn	
35 40 45	
ccc tgg ctg gga aac atc aaa gga ata att gga aag gtc ggg act ggg	192
Pro Trp Leu Gly Asn Ile Lys Gly Ile Ile Gly Lys Val Gly Thr Gly	
50 55 60	
gag gga gct ccc ccg gcg aag aga ccc cga aca gat cgg atg gag gtc	240
Glu Gly Ala Pro Pro Ala Lys Arg Pro Arg Thr Asp Arg Met Glu Val	
65 70 75 80	
gac tcc gga cca ggg aaa aag agt aac aag gga ggc ttc acc gac gag	288
Asp Ser Gly Pro Gly Lys Lys Ser Asn Lys Gly Gly Phe Thr Asp Glu	
85 90 95	
gag aga cgg gct cac cga cga agg aag gcc ctc gag aac aag aag aag	336
Glu Arg Arg Ala His Arg Arg Arg Lys Ala Leu Glu Asn Lys Lys Lys	
100 105 110	
cag ctg tcc gca gga gga aag tcc ctc agc aag gag gaa gaa gag gaa	384
Gln Leu Ser Ala Gly Gly Lys Ser Leu Ser Lys Glu Glu Glu Glu Glu	
115 120 125	
ctc agg aag ttg acc gcc gat gac gag caa aga gca aga aga ata gcc	432
Leu Arg Lys Leu Thr Ala Asp Asp Glu Gln Arg Ala Arg Arg Ile Ala	
130 135 140	
ggc ccg cgg gtt ggg gat gtg aac cct ccc gag gga cca ccg agg gga	480
Gly Pro Arg Val Gly Asp Val Asn Pro Pro Glu Gly Pro Pro Arg Gly	
145 150 155 160	
gcg ccc ggg ggc ggc ttc gtt cct caa ctg cta ggt gtc ccg gag tcc	528
Ala Pro Gly Gly Gly Phe Val Pro Gln Leu Leu Gly Val Pro Glu Ser	
165 170 175	
ccc ttc tcc aga acg gga gag gga ctc gac atc cga gga gac cga cag	576
Pro Phe Ser Arg Thr Gly Glu Gly Leu Asp Ile Arg Gly Asp Arg Gln	
180 185 190	
ttt cca tgg gga cct agt ccg acc ccc cct ccc cgc ctt cca ctc ctc	624
Phe Pro Trp Gly Pro Ser Pro Thr Pro Pro Pro Arg Leu Pro Leu Leu	
195 200 205	
gag tgt acc ccc caa taa	642
Glu Cys Thr Pro Gln	
210	

<211> 213

<212> PRT

<213> virus de la hepatitis D

5 <400> 3

```

Met Ser His Ala Asp Thr Lys Arg Ser Arg Lys Gly Arg Glu Glu Thr
 1           5           10           15

Leu Ser Lys Trp Xaa Lys Ala Arg Glu Asp Ala Glu Glu Leu Glu Arg
      20           25           30

Arg Leu Arg Lys Thr Lys Lys Thr Ile Lys Lys Leu Glu Asp Asp Asn
      35           40           45

Pro Trp Leu Gly Asn Ile Lys Gly Ile Ile Gly Lys Val Gly Thr Gly
      50           55           60

Glu Gly Ala Pro Pro Ala Lys Arg Pro Arg Thr Asp Arg Met Glu Val
      65           70           75           80

Asp Ser Gly Pro Gly Lys Lys Ser Asn Lys Gly Gly Phe Thr Asp Glu
      85           90           95

Glu Arg Arg Ala His Arg Arg Arg Lys Ala Leu Glu Asn Lys Lys Lys
      100           105           110

Gln Leu Ser Ala Gly Gly Lys Ser Leu Ser Lys Glu Glu Glu Glu Glu
      115           120           125

Leu Arg Lys Leu Thr Ala Asp Asp Glu Gln Arg Ala Arg Arg Ile Ala
      130           135           140

Gly Pro Arg Val Gly Asp Val Asn Pro Pro Glu Gly Pro Pro Arg Gly
      145           150           155           160

Ala Pro Gly Gly Gly Phe Val Pro Gln Leu Leu Gly Val Pro Glu Ser
      165           170           175

Pro Phe Ser Arg Thr Gly Glu Gly Leu Asp Ile Arg Gly Asp Arg Gln
      180           185           190

Phe Pro Trp Gly Pro Ser Pro Thr Pro Pro Pro Arg Leu Pro Leu Leu
      195           200           205

Glu Cys Thr Pro Gln
      210

```

<210> 4

<211> 585

<212> ADN

10 <213> virus de la hepatitis D

<220>

<221> CDS

<222> (1)..(585)

<223> ADNc de sHD

<400> 4

```

atg agc cac gcc gat acg aag aga tcc cgg aag gga agg gag gag acc 48
Met Ser His Ala Asp Thr Lys Arg Ser Arg Lys Gly Arg Glu Glu Thr
  1           5           10           15

ctt tcg aag tgg gwc aaa gca agg gaa gac gcg gaa gaa ctc gag agg 96
Leu Ser Lys Trp Xaa Lys Ala Arg Glu Asp Ala Glu Glu Leu Glu Arg
          20           25           30

agg ctc cgg aag aca aag aag acc atc aag aaa ctt gag gac gac aat 144
Arg Leu Arg Lys Thr Lys Lys Thr Ile Lys Lys Leu Glu Asp Asp Asn
          35           40           45

ccc tgg ctg gga aac atc aaa gga ata att gga aag gtc ggg act ggg 192
Pro Trp Leu Gly Asn Ile Lys Gly Ile Ile Gly Lys Val Gly Thr Gly
          50           55           60

gag gga gct ccc ccg gcg aag aga ccc cga aca gat cgg atg gag gtc 240
Glu Gly Ala Pro Pro Ala Lys Arg Pro Arg Thr Asp Arg Met Glu Val
          65           70           75           80

gac tcc gga cca ggg aaa aag agt aac aag gga ggc ttc acc gac gag 288
Asp Ser Gly Pro Gly Lys Lys Ser Asn Lys Gly Gly Phe Thr Asp Glu
          85           90           95

gag aga cgg gct cac cga cga agg aag gcc ctc gag aac aag aag aag 336
Glu Arg Arg Ala His Arg Arg Arg Lys Ala Leu Glu Asn Lys Lys Lys
          100           105           110

cag ctg tcc gca gga gga aag tcc ctc agc aag gag gaa gaa gag gaa 384
Gln Leu Ser Ala Gly Gly Lys Ser Leu Ser Lys Glu Glu Glu Glu
          115           120           125

ctc agg aag ttg acc gcc gat gac gag caa aga gca aga aga ata gcc 432
Leu Arg Lys Leu Thr Ala Asp Asp Glu Gln Arg Ala Arg Arg Ile Ala
          130           135           140

ggc ccg cgg gtt ggg gat gtg aac cct ccc gag gga cca ccg agg gga 480
Gly Pro Arg Val Gly Asp Val Asn Pro Pro Glu Gly Pro Pro Arg Gly
          145           150           155           160

gcg ccc ggg ggc ggc ttc gtt cct caa ctg cta ggt gtc ccg gag tcc 528
Ala Pro Gly Gly Gly Phe Val Pro Gln Leu Leu Gly Val Pro Glu Ser
          165           170           175

ccc ttc tcc aga acg gga gag gga ctc gac atc cga gga gac cga cag 576
Pro Phe Ser Arg Thr Gly Glu Gly Leu Asp Ile Arg Gly Asp Arg Gln
          180           185           190

ttt cca tag 585
Phe Pro
          195

```

<210> 5

<211> 194

<212> PRT

<213> virus de la hepatitis D

5

<400> 5

```

Met Ser His Ala Asp Thr Lys Arg Ser Arg Lys Gly Arg Glu Glu Thr
 1           5           10           15
Leu Ser Lys Trp Xaa Lys Ala Arg Glu Asp Ala Glu Glu Leu Glu Arg
      20           25           30
Arg Leu Arg Lys Thr Lys Lys Thr Ile Lys Lys Leu Glu Asp Asp Asn
      35           40           45
Pro Trp Leu Gly Asn Ile Lys Gly Ile Ile Gly Lys Val Gly Thr Gly
      50           55           60
Glu Gly Ala Pro Pro Ala Lys Arg Pro Arg Thr Asp Arg Met Glu Val
      65           70           75           80
Asp Ser Gly Pro Gly Lys Lys Ser Asn Lys Gly Gly Phe Thr Asp Glu
      85           90           95
Glu Arg Arg Ala His Arg Arg Arg Lys Ala Leu Glu Asn Lys Lys Lys
      100           105           110
Gln Leu Ser Ala Gly Gly Lys Ser Leu Ser Lys Glu Glu Glu Glu
      115           120           125
Leu Arg Lys Leu Thr Ala Asp Asp Glu Gln Arg Ala Arg Arg Ile Ala
      130           135           140
Gly Pro Arg Val Gly Asp Val Asn Pro Pro Glu Gly Pro Pro Arg Gly
      145           150           155           160
Ala Pro Gly Gly Gly Phe Val Pro Gln Leu Leu Gly Val Pro Glu Ser
      165           170           175
Pro Phe Ser Arg Thr Gly Glu Gly Leu Asp Ile Arg Gly Asp Arg Gln
      180           185           190
Phe Pro

```

10 <210> 6

<211> 1697

<212> ADN

<213> virus de la hepatitis D

15 <220>

<221> cepa

<222> (1)..(1697)

<223> secuencia genómica completa de dFr47

<400> 6

5

```

atggggccacg yagccgacaa ggagccggag gaggaggggt tggaggtccc gaggaggact 60
cyaagcaaag aagggggagga atctcgaaaa agaactccca agaagaaccg agaacaaccy 120
caagaagagg aggtttccca aacgcyggtg gagcaccggg gaacaaggaa acaggagaag 180
gatcggtaga aaagagcgag cctcccgatc cgagggggccc tggacacctc agctctggag 240
tcctccgacc cgtgaggttg aggttctccg cgcaggggnag gcagccaccg ggagaaaacg 300
gagaaaccac ctccagagga ccccttcagc gaacagagaa gcgcactggg acgtcaggat 360
gagtcctatg cgataggagg ggatgctagg agtcggcagc gaccgaagcg aggaaagaaa 420
gtaaagarag caacggggct agcgagtggg tgttccgcct ccccgagtc ccgagtgagg 480
cttatcccg ggaaccggc caagcccca aaatgtctgg tcccacggg ctccacttcc 540
aaaaggagga gcggggggac ttggaacgtg aggggaccag ttcagtccgt ggganttrcc 600
ccggacctcc ggttctccac actccttccc ccctgcgggc tcccagtaa acggtaaaga 660
cccacctcg gggtcggacg tcccacctt tcttaccttg tggccggcat ggtcccagcc 720
tcctcgctgg cgccggctgg gcaacattcc grggggaccg tccctcggtg atggcgaatg 780
ggccccagaa ctctctctag gttcccagag agaaccgaga gaaaactggc tctcccttag 840
ccatccgagt ggacgttctg tctccttcg gatgccagag tcggaccgag gggaggtgga 900
gatgccatgc cgaccgaag aggaaagaag gactcggacg cgaactcgtg agtggaaaact 960
tattccttta ctggggagta cactcgagga gtggaaggcg gggaggccgg ggtcccgggy 1020
ytaccacag gtactgctgg ttcccctca cgtccagtc ctccccgct ctggagaagg 1080
gagactcrgg aacacctagc atttgaggga caaagccgcc cccgggcgct cccctcgag 1140
atgttcggg tgggttcaca tcccacccc gcgggcccgc tactcttct tctcttctct 1200
cgtcttctc ggtcaacttc cggagttcct ctctctctc ctctgctragt gactttctct 1260
ccgcgctcag ttgtttcttc ttgttctcga gggccttct tcttcggttg tcccgctctct 1320
ccttgctcgt gaagtctccc gcgcgagtc tcttcttagg tccggagtcg acctccatcy 1380
gatccgctcc ggccctcttc gccgggggag ctccctcccc gtccttccct tttcttatta 1440
ttccgaggat gttccccagc cagggattct catcctcgag tytcttgcg gttctcttg 1500
ccttcggag ccttttctcg aggtcctccg cgtcttttct tgcttgatc catttctcga 1560
ggacatctc ccttctctcc ctccggggtt tcctctgttc ggactggctc atcctcgag 1620
agggggcgac ggcgtctgtt ctctattctt ttttctctaa aagaggagac agcttgccgc 1680
ccgccccag ccccggg 1697

```

<210> 7

<211> 645

10 <212> ADN

<213> virus de la hepatitis D

<220>

<221> CDS

15 <222> (1)..(645)

<223> ADNc de LHD

<400> 7

atg agc cag tcc gaa cag agg aaa ccc cgg agg gga gga agg gag gat	48
Met Ser Gln Ser Glu Gln Arg Lys Pro Arg Arg Gly Gly Arg Glu Asp	
1 5 10 15	
gtc ctc gag aaa tgg atc caa gca aga aaa gac gcg gag gac ctc gag	96
Val Leu Glu Lys Trp Ile Gln Ala Arg Lys Asp Ala Glu Asp Leu Glu	
20 25 30	
aaa agg ctc cgg aag gcc aag aga acc gcc aag ara ctc gag gat gag	144
Lys Arg Leu Arg Lys Ala Lys Arg Thr Ala Lys Xaa Leu Glu Asp Glu	
35 40 45	
aat ccc tgg ctg ggg aac atc ctc gga ata ata aga aaa ggg aag gac	192
Asn Pro Trp Leu Gly Asn Ile Leu Gly Ile Ile Arg Lys Gly Lys Asp	
50 55 60	
ggg gag gga gct ccc ccg gcg aag agg gcc cgg acg gat crg atg gag	240
Gly Glu Gly Ala Pro Pro Ala Lys Arg Ala Arg Thr Asp Xaa Met Glu	
65 70 75 80	
gtc gac tcc gga cct agg aag agg act cgc gcg gga gac ttc acc gac	288
Val Asp Ser Gly Pro Arg Lys Arg Thr Arg Ala Gly Asp Phe Thr Asp	
85 90 95	
aag gag aga cgg gac cac cga aga agg aag gcc ctc gag aac aag aag	336
Lys Glu Arg Arg Asp His Arg Arg Arg Lys Ala Leu Glu Asn Lys Lys	
100 105 110	
aaa caa ctg agc gcg gga gga aag tca cty agc aag gag gag gaa gag	384
Lys Gln Leu Ser Ala Gly Gly Lys Ser Xaa Ser Lys Glu Glu Glu Glu	
115 120 125	
gaa ctc cgg aag ttg acc gag gaa gac gag aga aga gaa aga aga gta	432
Glu Leu Arg Lys Leu Thr Glu Glu Asp Glu Arg Arg Glu Arg Arg Val	
130 135 140	
gcc ggc ccg cgg gtt ggg gat gtg aac cca ccc gga aca tct ccg agg	480
Ala Gly Pro Arg Val Gly Asp Val Asn Pro Pro Gly Thr Ser Pro Arg	
145 150 155 160	
gga gcg ccc ggg ggc ggc ttt gtc cct caa atg cta ggt gtt ccy gag	528
Gly Ala Pro Gly Gly Gly Phe Val Pro Gln Met Leu Gly Val Xaa Glu	
165 170 175	
tct ccc ttc tcc agg acg ggg gag gga ctg gac gtg agg ggg aac cag	576
Ser Pro Phe Ser Arg Thr Gly Glu Gly Leu Asp Val Arg Gly Asn Gln	
180 185 190	
cag tac ccg tgg gtg grc ccg gga ccc cgg cct ccc cgc ctt cca ctc	624
Gln Tyr Pro Trp Val Xaa Pro Gly Pro Arg Pro Pro Arg Leu Pro Leu	
195 200 205	
ctc gag tgt act ccc cag taa	645
Leu Glu Cys Thr Pro Gln	
210 215	

<210> 8

5 <211> 214

<212> PRT

<213> virus de la hepatitis D

<400> 8

Met Ser Gln Ser Glu Gln Arg Lys Pro Arg Arg Gly Gly Arg Glu Asp
 1 5 10 15

 Val Leu Glu Lys Trp Ile Gln Ala Arg Lys Asp Ala Glu Asp Leu Glu
 20 25 30
 Lys Arg Leu Arg Lys Ala Lys Arg Thr Ala Lys Xaa Leu Glu Asp Glu
 35 40 45
 Asn Pro Trp Leu Gly Asn Ile Leu Gly Ile Ile Arg Lys Gly Lys Asp
 50 55 60
 Gly Glu Gly Ala Pro Pro Ala Lys Arg Ala Arg Thr Asp Xaa Met Glu
 65 70 75 80
 Val Asp Ser Gly Pro Arg Lys Arg Thr Arg Ala Gly Asp Phe Thr Asp
 85 90 95
 Lys Glu Arg Arg Asp His Arg Arg Arg Lys Ala Leu Glu Asn Lys Lys
 100 105 110
 Lys Gln Leu Ser Ala Gly Gly Lys Ser Xaa Ser Lys Glu Glu Glu Glu
 115 120 125
 Glu Leu Arg Lys Leu Thr Glu Glu Asp Glu Arg Arg Glu Arg Arg Val
 130 135 140
 Ala Gly Pro Arg Val Gly Asp Val Asn Pro Pro Gly Thr Ser Pro Arg
 145 150 155 160
 Gly Ala Pro Gly Gly Gly Phe Val Pro Gln Met Leu Gly Val Xaa Glu
 165 170 175
 Ser Pro Phe Ser Arg Thr Gly Glu Gly Leu Asp Val Arg Gly Asn Gln
 180 185 190
 Gln Tyr Pro Trp Val Xaa Pro Gly Pro Arg Pro Pro Arg Leu Pro Leu
 195 200 205
 Leu Glu Cys Thr Pro Gln
 210

5

<210> 9

<211> 588

<212> ADN

10 <213> virus de la hepatitis D

<220>

<221> CDS

<222> (1)..(588)

<223> ADNc de sHD

5 <400> 9

atg agc cag tcc gaa cag agg aaa ccc cgg agg gga gga agg gag gat 48
 Met Ser Gln Ser Glu Gln Arg Lys Pro Arg Arg Gly Gly Arg Glu Asp
 1 5 10 15

gtc ctc gag aaa tgg atc caa gca aga aaa gac gcg gag gac ctc gag 96
 Val Leu Glu Lys Trp Ile Gln Ala Arg Lys Asp Ala Glu Asp Leu Glu
 20 25 30

aaa agg ctc cgg aag gcc aag aga acc gcc aag ara ctc gag gat gag 144
 Lys Arg Leu Arg Lys Ala Lys Arg Thr Ala Lys Xaa Leu Glu Asp Glu
 35 40 45

aat ccc tgg ctg ggg aac atc ctc gga ata ata aga aaa ggg aag gac 192
 Asn Pro Trp Leu Gly Asn Ile Leu Gly Ile Ile Arg Lys Gly Lys Asp
 50 55 60

ggg gag gga gct ccc ccg gcg aag agg gcc cgg acg gat crg atg gag 240
 Gly Glu Gly Ala Pro Pro Ala Lys Arg Ala Arg Thr Asp Xaa Met Glu
 65 70 75 80

gtc gac tcc gga cct agg aag agg act cgc gcg gga gac ttc acc gac 288
 Val Asp Ser Gly Pro Arg Lys Arg Thr Arg Ala Gly Asp Phe Thr Asp
 85 90 95

aag gag aga cgg gac cac cga aga agg aag gcc ctc gag aac aag aag 336
 Lys Glu Arg Arg Asp His Arg Arg Arg Lys Ala Leu Glu Asn Lys Lys
 100 105 110

aaa caa ctg agc gcg gga gga aag tca cty agc aag gag gag gaa gag 384
 Lys Gln Leu Ser Ala Gly Gly Lys Ser Xaa Ser Lys Glu Glu Glu Glu
 115 120 125

gaa ctc cgg aag ttg acc gag gaa gac gag aga aga gaa aga aga gta 432
 Glu Leu Arg Lys Leu Thr Glu Glu Asp Glu Arg Arg Glu Arg Arg Val
 130 135 140

gcc ggc ccg cgg gtt ggg gat gtg aac cca ccc gga aca tct ccg agg 480
 Ala Gly Pro Arg Val Gly Asp Val Asn Pro Pro Gly Thr Ser Pro Arg
 145 150 155 160

gga gcg ccc ggg ggc ggc ttt gtc cct caa atg cta ggt gtt ccy gag 528
 Gly Ala Pro Gly Gly Gly Phe Val Pro Gln Met Leu Gly Val Xaa Glu
 165 170 175

tct ccc ttc tcc agg acg ggg gag gga ctg gac gtg agg ggg aac cag 576
 Ser Pro Phe Ser Arg Thr Gly Glu Gly Leu Asp Val Arg Gly Asn Gln
 180 185 190

cag tac ccg tga 588
 Gln Tyr Pro
 195

<210> 10

<211> 195

5 <212> PRT

<213> virus de la hepatitis D

Met Ser Gln Ser Glu Gln Arg Lys Pro Arg Arg Gly Gly Arg Glu Asp
1 5 10 15

Gln Tyr Pro
195

<223> secuencia genómica completa de dFr73

<400> 11

```

atgggcctcg agccgacgaa gagccggaga gggaggggtc ggggaycccg aggaggraac 60
caagcaaaga aggggagact cccgaggraa grtatctcct aagaagagtc gagaarattc 120
tcaagaggag gaggwatccc ctcaaagctg acggaactcc gggaacaagg aaayaggaga 180
aggaacggta gaaaagagcg agcctcccga tccgaggggc ccggacacct cggctctgga 240
gtcctccgac ctgtgagggtc gagaacctcc gcgcagggag acagccacca ggagaaggca 300
gaaaaaccac ctccagagga ccccttcagc gaacagagaa gcactctgag gcgtcrggag 360
agnntccata gcgataggag gagatgctag gagtgggcgg cgaccraagc gaggatcgaa 420
agtaaagaaa gcaacggggc tagcaggtgg atgttccgcc tccccagggt cccgagtggag 480
gcttatcccc gggaacccgg ctaagccccg aaatgtctgg ttccgccggc tccacttcca 540

```

```

aaaggaggag cgggggggact tggarcgyaa ggggccagtt cagtccgtgg ggwttacccc 600
ggagctcccc ttctccacac tccttcccc cgcgggaccc ccaataatnt ggtaagaacc 660
caccactggg ggtccgacgy tcctaccctt tcttacctgt ggccggcatg gtccagacct 720
cctcgctggc gccggtggg caacattccg aggggacggc ctctcggta atggcgaatg 780
ggaccagaw ctctctctag gttcccagag agaaccgaga gaaaactggc tctcccttag 840
ccatccgagt ggacgtctgt cctccttcgg atgcccaggt cggaccgcgg ggaggtggag 900
atgccatgcc gacccgaaga ggaaagaagg actcggacgc gaaccctga gtggaacctt 960
attcctttat tggggagtac actcaggag tggaaggcgg ggagacgggg ctcccgggct 1020
taccyacggg aactgctggt tcccccttat gtccagtcct tctcccgctc tgggtaaggr 1080
agactcggga acgttcagca tttgaggggac aaagccggcc ccgggcgctc ccctcggagg 1140
tcctgtggtt ggggtcacat ccccaacccg cgggcccggc actcttcttt cccttctctc 1200
gtcttcytcg gtcaacctcc cgagttcctc ttcttcctcc ttgctgagtg actttccycc 1260
cgcggttcagc tgtttcttct tgttctcgag ggccttcctt ctccggtggg cctgtctctc 1320
cttgctgggt aagtctcccc cgcgagatct cttcctaggt ccggaatcga cctccatccg 1380
atccgtccgg gccctcttcg ccgggggagc tccctccccg tccttccctt ttcttattat 1440
tccgaggacg ttccccagcc agggattttc gtcctcgagt ctcttgatgg ttctcttggt 1500
cttcggagc cttytctcga ggtctcgcg rtcttttctt gcttggaycc atttctcgag 1560
grtctctcc ctctctccc tccgggattt ctctgttcg gactggctca tctcggata 1620
gggggcgacg gcgtctgttc tctattttct tcgttctaag aagaggagay agcgtgccgc 1680
ccgccccag ccccggg T697

```

5

<210> 12

<211> 645

<212> ADN

<213> virus de la hepatitis D

10

<220>

<221> CDS

<222> (1)..(695)

<223> ADNc de LHD

15

<400> 12

atg agc cag tcc gaa cag aag aaa tcc cgg agg gga gga agg gag gag	48
Met Ser Gln Ser Glu Gln Lys Lys Ser Arg Arg Gly Gly Arg Glu Glu	
1 5 10 15	
ayc ctc gag aaa tgg rtc caa gca aga aaa gay gcg gas gac ctc gag	96
Xaa Leu Glu Lys Trp Xaa Gln Ala Arg Lys Asp Ala Xaa Asp Leu Glu	
20 25 30	
ara agg ctc cgg aag acc aag aga acc atc aag aga ctc gag gac gaa	144
Xaa Arg Leu Arg Lys Thr Lys Arg Thr Ile Lys Arg Leu Glu Asp Glu	
35 40 45	
aat ccc tgg ctg ggg aac gtc ctc gga ata ata aga aaa ggg aag gac	192
Asn Pro Trp Leu Gly Asn Val Leu Gly Ile Ile Arg Lys Gly Lys Asp	
50 55 60	
ggg gag gga gct ccc ccg gcg aag agg gcc cgg acg gat cgg atg gag	240
Gly Glu Gly Ala Pro Pro Ala Lys Arg Ala Arg Thr Asp Arg Met Glu	
65 70 75 80	
gtc gat tcc gga cct agg aag aga tct cgc gcg gga gac ttc acc gac	288
Val Asp Ser Gly Pro Arg Lys Arg Ser Arg Ala Gly Asp Phe Thr Asp	
85 90 95	
aag gag aga cag gac cac cga aga agg aag gcc ctc gag aac aag aag	336
Lys Glu Arg Gln Asp His Arg Arg Arg Lys Ala Leu Glu Asn Lys Lys	
100 105 110	
aaa cag ctg aac gcg ggr gga aag tca ctc agc aag gag gaa gaa gag	384
Lys Gln Leu Asn Ala Xaa Gly Lys Ser Leu Ser Lys Glu Glu Glu Glu	
115 120 125	
gaa ctc ggg agg ttg acc gar gaa gac gag aga agg gaa aga aga gta	432
Glu Leu Gly Arg Leu Thr Glu Glu Asp Glu Arg Arg Glu Arg Arg Val	
130 135 140	
gcc ggc ccg cgg gtt ggg gat gtg aac cca cca cag gga cct ccg agg	480
Ala Gly Pro Arg Val Gly Asp Val Asn Pro Pro Gln Gly Pro Pro Arg	
145 150 155 160	
gga gcg ccc ggg ggc ggc ttt gtc cct caa atg ctg aac gtt ccc gag	528
Gly Ala Pro Gly Gly Gly Phe Val Pro Gln Met Leu Asn Val Pro Glu	
165 170 175	
tct ycc ttc acc agg acg gga gag gga ctg gac ata agg ggg aac cag	576
Ser Xaa Phe Thr Arg Thr Gly Glu Gly Leu Asp Ile Arg Gly Asn Gln	
180 185 190	
cag ttc ccg tgg gta agc ccg gga gcc ccg tct ccc cgc ctt cca ctc	624
Gln Phe Pro Trp Val Ser Pro Gly Ala Pro Ser Pro Arg Leu Pro Leu	
195 200 205	
ctc gag tgt act ccc caa taa	645
Leu Glu Cys Thr Pro Gln	
210 215	

<211> 214

<212> PRT

<213> virus de la hepatitis D

5

<400> 13

Met Ser Gln Ser Glu Gln Lys Lys Ser Arg Arg Gly Gly Arg Glu Glu
 1 5 10 15

Xaa Leu Glu Lys Trp Xaa Gln Ala Arg Lys Asp Ala Xaa Asp Leu Glu
 20 25 30

Xaa Arg Leu Arg Lys Thr Lys Arg Thr Ile Lys Arg Leu Glu Asp Glu
 35 40 45

Asn Pro Trp Leu Gly Asn Val Leu Gly Ile Ile Arg Lys Gly Lys Asp
 50 55 60

Gly Glu Gly Ala Pro Pro Ala Lys Arg Ala Arg Thr Asp Arg Met Glu
 65 70 75 80

Val Asp Ser Gly Pro Arg Lys Arg Ser Arg Ala Gly Asp Phe Thr Asp
 85 90 95

Lys Glu Arg Gln Asp His Arg Arg Arg Lys Ala Leu Glu Asn Lys Lys
 100 105 110

Lys Gln Leu Asn Ala Xaa Gly Lys Ser Leu Ser Lys Glu Glu Glu Glu
 115 120 125

Glu Leu Gly Arg Leu Thr Glu Glu Asp Glu Arg Arg Glu Arg Arg Val
 130 135 140

Ala Gly Pro Arg Val Gly Asp Val Asn Pro Pro Gln Gly Pro Pro Arg
 145 150 155 160

Gly Ala Pro Gly Gly Gly Phe Val Pro Gln Met Leu Asn Val Pro Glu
 165 170 175

Ser Xaa Phe Thr Arg Thr Gly Glu Gly Leu Asp Ile Arg Gly Asn Gln
 180 185 190

Gln Phe Pro Trp Val Ser Pro Gly Ala Pro Ser Pro Arg Leu Pro Leu
 195 200 205

Leu Glu Cys Thr Pro Gln
 210

10 <210> 14

<211> 588

<212> ADN

<213> virus de la hepatitis D

<220>

<221> CDS

5 <222> (1)..(588)

<223> ADNc de sHD

<400> 14


```

atg agc cag tcc gaa cag aag aaa tcc cgg agg gga gga agg gag gag 48
Met Ser Gln Ser Glu Gln Lys Lys Ser Arg Arg Gly Gly Arg Glu Glu
1 5 10 15

ayc ctc gag aaa tgg rtc caa gca aga aaa gay gcg gas gac ctc gag 96
Xaa Leu Glu Lys Trp Xaa Gln Ala Arg Lys Asp Ala Xaa Asp Leu Glu
20 25 30

ara agg ctc cgg aag acc aag aga acc atc aag aga ctc gag gac gaa 144
Xaa Arg Leu Arg Lys Thr Lys Arg Thr Ile Lys Arg Leu Glu Asp Glu
35 40 45

aat ccc tgg ctg ggg aac gtc ctc gga ata ata aga aaa ggg aag gac 192
Asn Pro Trp Leu Gly Asn Val Leu Gly Ile Ile Arg Lys Gly Lys Asp
50 55 60

ggg gag gga gct ccc ccg gcg aag agg gcc cgg acg gat cgg atg gag 240
Gly Glu Gly Ala Pro Pro Ala Lys Arg Ala Arg Thr Asp Arg Met Glu
65 70 75 80

gtc gat tcc gga cct agg aag aga tct cgc gcg gga gac ttc acc gac 288
Val Asp Ser Gly Pro Arg Lys Arg Ser Arg Ala Gly Asp Phe Thr Asp
85 90 95

aag gag aga cag gac cac cga aga agg aag gcc ctc gag aac aag aag 336
Lys Glu Arg Gln Asp His Arg Arg Arg Lys Ala Leu Glu Asn Lys Lys
100 105 110

aaa cag ctg aac gcg ggr gga aag tca ctc agc aag gag gaa gaa gag 384
Lys Gln Leu Asn Ala Xaa Gly Lys Ser Leu Ser Lys Glu Glu Glu Glu
115 120 125

gaa ctc ggg agg ttg acc gar gaa gac gag aga agg gaa aga aga gta 432
Glu Leu Gly Arg Leu Thr Glu Glu Asp Glu Arg Arg Glu Arg Arg Val
130 135 140

gcc gcc ccg cgg gtt ggg gat gtg aac cca cca cag gga cct ccg agg 480
Ala Gly Pro Arg Val Gly Asp Val Asn Pro Pro Gln Gly Pro Pro Arg
145 150 155 160

gga gcg ccc ggg ggc ggc ttt gtc cct caa atg ctg aac gtt ccc gag 528
Gly Ala Pro Gly Gly Gly Phe Val Pro Gln Met Leu Asn Val Pro Glu
165 170 175

tct ycc ttc acc agg acg gga gag gga ctg gac ata agg ggg aac cag 576
Ser Xaa Phe Thr Arg Thr Gly Glu Gly Leu Asp Ile Arg Gly Asn Gln
180 185 190

cag ttc ccg tag 588
Gln Phe Pro
195

```

<210> 15

<211> 195

5 <212> PRT

<213> virus de la hepatitis D

<400> 15

```

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

```

```

Lys Glu Arg Gln Asp His Arg Arg Arg Lys Ala Leu Glu Asn Lys Lys
          100           105           110
Lys Gln Leu Asn Ala Xaa Gly Lys Ser Leu Ser Lys Glu Glu Glu Glu
          115           120           125
Glu Leu Gly Arg Leu Thr Glu Glu Asp Glu Arg Arg Glu Arg Arg Val
          130           135           140
Ala Gly Pro Arg Val Gly Asp Val Asn Pro Pro Gln Gly Pro Pro Arg
          145           150           155           160
Gly Ala Pro Gly Gly Gly Phe Val Pro Gln Met Leu Asn Val Pro Glu
          165           170           175
Ser Xaa Phe Thr Arg Thr Gly Glu Gly Leu Asp Ile Arg Gly Asn Gln
          180           185           190
Gln Phe Pro
          195

```

5

<210> 16

<211> 1697

<212> ADN

10 <213> virus de la hepatitis D

<220>

<221> cepa

<222> (1) .. (1697)

<223> secuencia genómica completa de dFr910

<400> 16

```

atggggcctcc gtgccgacga agaateggat cgataggaag tgggaggccc gagaggggct 60
ccaagcaaag agaggaggaa tcccggggaa aaattctccc aagaagagtc gagagtaatc 120
tcaagaagag gagatttccc agacgccggg ggagcaccag ggaacaagga aacaggagag 180
gaacggtaga aaagagcgag cctcccgatc cgagggggcc ctggaccctt cagctctgga 240
gtcctccgac ccgtgaggtc gaggatcccc gcgcaggagg acagccaccc ggagaagaca 300
gagaaatcac ctccagagaa accctctcag cgaacaaaga ggcgctctga ggcgttgggg 360
cgagtccata gcgatgggag gggatgctag gattcggcag cgaccgaagc gaggagaaaa 420
gtaaagaaag caacggggct agcagggtga tgttccgctt ccccgagggt cccgagttag 480
gcttatcccg gggaactcgg ctaagccccg aaatgtctag ttccactgtc tccacttcca 540
aaggaggagg cggttgacc tggaaatga ggggaccagt caatccgtgg gattgccccg 600
gacctccggy tctccacact ccttcccccc tgcgggcttc ccagtaaaat gaaagaaacc 660
cacccactgg ggtccgacgy cctccccctt tcttacctgt ggccggcatg gtcccagcct 720
ccycgctggc gccggctggg caacattccg aggggaccgt ccctcggtaa tggcgaatgg 780
gaccagaac tctctccagg ttcccagaga gaaccgagag aaaactggct ctcccttagc 840
catccgagtg gacgttyygt cctccttcgg atgccaggt cggaccgcgg ggaggtggag 900
aygccatgcc gaccogaaga aggaagaag gactcggacg cgaaccctgt agtggaactc 960
tttatctttt actggggagt aactcagag agtggaaggc ggggaggccg ggggtccggg 1020
tctaccacag ggaattgctg gtttcccctt acgtccagtc cacttcccgt cctggagaag 1080
ggagactcgg gaacatctag catttgaggg acaaagccgc ccccgggcgc tcccctcgga 1140
ggtccctcgg gtgggttcac atccccaaac cgcgggcgg ctactcttct tccccttctc 1200
tcgtcttcct cggccaactt ccggagttec tcttctctct ccctgctgag tgactttccc 1260
ccgcgcgtca gctgattctt cttgttctcg agggctttcc ttcttcgggt atcctttctc 1320
tcctggtcgg tgaagtctcc cgctcgggtc ctcttcttag gtccggagtc gacctccatc 1380
tgatccgtcc gggccctttt cgcgggggga gtcctctccc cgtccttccc ttttcttatt 1440

```

```

attccaagga tgttccccag ccagggattt tcatcctoga gtttcttgat ggttctcttg 1500
gtcttccgga gcyttttctc gagatcctcc gactcttttc ttgcttgac ccacttctcg 1560
aggatctcct cccttctctc tctccgagct ttcttcgatt cggactggct catcctcgga 1620
tagggggaga cgcttctgt totcttattc ttcttttcg aagaggagac agcttgccgc 1680
ccctccccca ccccggg 1697

```

5

<210> 17

<211> 645

<212> ADN

10 <213> virus de la hepatitis D

<220>

<221> CDS

<222> (1)..(645)

15 <223> ADNc de LHD

<400> 17

atg agc cag tcc gaa tcg aag aaa gct cgg aga gga gga agg gag gag	48
Met Ser Gln Ser Glu Ser Lys Lys Ala Arg Arg Gly Gly Arg Glu Glu	
1 5 10 15	
atc ctc gag aag tgg gtg caa gca aga aaa gac tcg gag gat ctc gag	96
Ile Leu Glu Lys Trp Val Gln Ala Arg Lys Asp Ser Glu Asp Leu Glu	
20 25 30	
aaa arg ctc cgg aag acc aag aga acc atc aag aaa ctc gag gat gaa	144
Lys Xaa Leu Arg Lys Thr Lys Arg Thr Ile Lys Lys Leu Glu Asp Glu	
35 40 45	
aat ccc tgg ctg ggg aac atc ctt gga ata ata aga aaa ggg aag gac	192
Asn Pro Trp Leu Gly Asn Ile Leu Gly Ile Ile Arg Lys Gly Lys Asp	
50 55 60	
ggg gag gga gct ccc ccg gcg aaa agg gcc cgg acg gat cag atg gag	240
Gly Glu Gly Ala Pro Pro Ala Lys Arg Ala Arg Thr Asp Gln Met Glu	
65 70 75 80	
gtc gac tcc gga cct agg aag agg acc cga gcg gga gac ttc acc gac	288
Val Asp Ser Gly Pro Arg Lys Arg Thr Arg Ala Gly Asp Phe Thr Asp	
85 90 95	
cag gag aga aag gat cac cga aga agg aaa gcc ctc gag aac aag aag	336
Gln Glu Arg Lys Asp His Arg Arg Arg Lys Ala Leu Glu Asn Lys Lys	
100 105 110	
aat cag ctg agc gcg ggg gga aag tca ctc agc agg gag gag gaa gag	384
Asn Gln Leu Ser Ala Gly Gly Lys Ser Leu Ser Arg Glu Glu Glu Glu	
115 120 125	
gaa ctc cgg aag ttg acc gag gaa gac gag aga agg gaa aga aga gta	432
Glu Leu Arg Lys Leu Thr Glu Glu Asp Glu Arg Arg Glu Arg Arg Val	
130 135 140	
gcc ggc ccg cgg gtt ggg gat gtg aac cca ccc gag gga cct ccg agg	480
Ala Gly Pro Arg Val Gly Asp Val Asn Pro Pro Glu Gly Pro Pro Arg	
145 150 155 160	
gga gcg ccc ggg ggc ggc ttt gtc cct caa atg cta gat gtt ccc gag	528
Gly Ala Pro Gly Gly Gly Phe Val Pro Gln Met Leu Asp Val Pro Glu	
165 170 175	
tct ccc ttc tcc agg acg gga agt gga ctg gac gta agg gga aac cag	576
Ser Pro Phe Ser Arg Thr Gly Ser Gly Leu Asp Val Arg Gly Asn Gln	
180 185 190	
caa ttc ccg tgg gtg gac ccg gga ccc cgg cct ccc cgc ctt cca ctc	624
Gln Phe Pro Trp Val Asp Pro Gly Pro Arg Pro Pro Arg Leu Pro Leu	
195 200 205	
ctc gag tgt act ccc cag taa	645
Leu Glu Cys Thr Pro Gln	
210 215	

<211> 214

<212> PRT

<213> virus de la hepatitis D

5 <400> 18

Met Ser Gln Ser Glu Ser Lys Lys Ala Arg Arg Gly Gly Arg Glu Glu
 1 5 10 15

Ile Leu Glu Lys Trp Val Gln Ala Arg Lys Asp Ser Glu Asp Leu Glu
 20 25 30

Lys Xaa Leu Arg Lys Thr Lys Arg Thr Ile Lys Lys Leu Glu Asp Glu
 35 40 45

Asn Pro Trp Leu Gly Asn Ile Leu Gly Ile Ile Arg Lys Gly Lys Asp
 50 55 60

Gly Glu Gly Ala Pro Pro Ala Lys Arg Ala Arg Thr Asp Gln Met Glu
 65 70 75 80

Val Asp Ser Gly Pro Arg Lys Arg Thr Arg Ala Gly Asp Phe Thr Asp
 85 90 95

Gln Glu Arg Lys Asp His Arg Arg Arg Lys Ala Leu Glu Asn Lys Lys
 100 105 110

Asn Gln Leu Ser Ala Gly Gly Lys Ser Leu Ser Arg Glu Glu Glu Glu
 115 120 125

Glu Leu Arg Lys Leu Thr Glu Glu Asp Glu Arg Arg Glu Arg Arg Val
 130 135 140

Ala Gly Pro Arg Val Gly Asp Val Asn Pro Pro Glu Gly Pro Pro Arg
 145 150 155 160

Gly Ala Pro Gly Gly Gly Phe Val Pro Gln Met Leu Asp Val Pro Glu
 165 170 175

Ser Pro Phe Ser Arg Thr Gly Ser Gly Leu Asp Val Arg Gly Asn Gln
 180 185 190

Gln Phe Pro Trp Val Asp Pro Gly Pro Arg Pro Pro Arg Leu Pro Leu
 195 200 205

Leu Glu Cys Thr Pro Gln
 210

<210> 19

10 <211> 588

<212> ADN

<213> virus de la hepatitis D

<220>

<221> CDS

<222> (1)..(588)

5 <223> ADNc de sHD

<400> 19

atg agc cag tcc gaa tcg aag aaa gct cgg aga gga gga agg gag gag	48
Met Ser Gln Ser Glu Ser Lys Lys Ala Arg Arg Gly Gly Arg Glu Glu	
1 5 10 15	
atc ctc gag aag tgg gtg caa gca aga aaa gac tcg gag gat ctc gag	96
Ile Leu Glu Lys Trp Val Gln Ala Arg Lys Asp Ser Glu Asp Leu Glu	
20 25 30	
aaa arg ctc cgg aag acc aag aga acc atc aag aaa ctc gag gat gaa	144
Lys Xaa Leu Arg Lys Thr Lys Arg Thr Ile Lys Lys Leu Glu Asp Glu	
35 40 45	
aat ccc tgg ctg ggg aac atc ctt gga ata ata aga aaa ggg aag gac	192
Asn Pro Trp Leu Gly Asn Ile Leu Gly Ile Ile Arg Lys Gly Lys Asp	
50 55 60	
ggg gag gga gct ccc ccg gcg aaa agg gcc cgg acg gat cag atg gag	240
Gly Glu Gly Ala Pro Pro Ala Lys Arg Ala Arg Thr Asp Gln Met Glu	
65 70 75 80	
gtc gac tcc gga cct agg aag agg acc cga gcg gga gac ttc acc gac	288
Val Asp Ser Gly Pro Arg Lys Arg Thr Arg Ala Gly Asp Phe Thr Asp	
85 90 95	
cag gag aga aag gat cac cga aga agg aaa gcc ctc gag aac aag aag	336
Gln Glu Arg Lys Asp His Arg Arg Lys Ala Leu Glu Asn Lys Lys	
100 105 110	
aat cag ctg agc gcg ggg gga aag tca ctc agc agg gag gag gaa gag	384
Asn Gln Leu Ser Ala Gly Gly Lys Ser Leu Ser Arg Glu Glu Glu Glu	
115 120 125	
gaa ctc cgg aag ttg acc gag gaa gac gag aga agg gaa aga aga gta	432
Glu Leu Arg Lys Leu Thr Glu Glu Asp Glu Arg Arg Glu Arg Arg Val	
130 135 140	
gcc ggc ccg cgg gtt ggg gat gtg aac cca ccc gag gga cct ccg agg	480
Ala Gly Pro Arg Val Gly Asp Val Asn Pro Pro Glu Gly Pro Pro Arg	
145 150 155 160	
gga gcg ccc ggg ggc ggc ttt gtc cct caa atg cta gat gtt ccc gag	528
Gly Ala Pro Gly Gly Gly Phe Val Pro Gln Met Leu Asp Val Pro Glu	
165 170 175	
tct ccc ttc tcc agg acg gga agt gga ctg gac gta agg gga aac cag	576
Ser Pro Phe Ser Arg Thr Gly Ser Gly Leu Asp Val Arg Gly Asn Gln	
180 185 190	
caa ttc ccg tga	588
Gln Phe Pro	
195	

<210> 20

<211> 195

<212> PRT

<213> virus de la hepatitis D

<400> 20

5

```

Met Ser Gln Ser Glu Ser Lys Lys Ala Arg Arg Gly Gly Arg Glu Glu
 1           5           10           15

Ile Leu Glu Lys Trp Val Gln Ala Arg Lys Asp Ser Glu Asp Leu Glu
      20           25           30

Lys Xaa Leu Arg Lys Thr Lys Arg Thr Ile Lys Lys Leu Glu Asp Glu
      35           40           45

Asn Pro Trp Leu Gly Asn Ile Leu Gly Ile Ile Arg Lys Gly Lys Asp
      50           55           60

Gly Glu Gly Ala Pro Pro Ala Lys Arg Ala Arg Thr Asp Gln Met Glu
      65           70           75           80

Val Asp Ser Gly Pro Arg Lys Arg Thr Arg Ala Gly Asp Phe Thr Asp
      85           90           95

Gln Glu Arg Lys Asp His Arg Arg Arg Lys Ala Leu Glu Asn Lys Lys
      100           105           110

Asn Gln Leu Ser Ala Gly Gly Lys Ser Leu Ser Arg Glu Glu Glu Glu
      115           120           125

Glu Leu Arg Lys Leu Thr Glu Glu Asp Glu Arg Arg Glu Arg Arg Val
      130           135           140

Ala Gly Pro Arg Val Gly Asp Val Asn Pro Pro Glu Gly Pro Pro Arg
      145           150           155           160

Gly Ala Pro Gly Gly Gly Phe Val Pro Gln Met Leu Asp Val Pro Glu
      165           170           175

Ser Pro Phe Ser Arg Thr Gly Ser Gly Leu Asp Val Arg Gly Asn Gln
      180           185           190

Gln Phe Pro
      195

```

<210> 21

<211> 1687

10 <212> ADN

<213> virus de la hepatitis D

<220>

<221> cepa

<222> (1)..(1687)

<223> secuencia genómica completa de dFr48

<400> 21

```

agtgggccac aagccggcgg agagatcgag gatttgggga ggagagggga accgaggggg 60
gttcgaactc aagaagagaa gatttttcgag gaaatactcc caagaagttc ccggaggggat 120
ccaagagacg gaygactttc cccyattggt tgggaaaagt cccggaacca ggaaatggga 180
aagagacggt agaaaagagc gagcctccca gttttcgagt agtcccggac ctatctaaat 240
tggagtcata cgaccgtatg gagaaaattt ccgacaagag gtgatccact gagggtttgc 300
ggaaaaatca cctccagagg accccttcag cgaacggaag agactggaag cccagagga 360
agaccatagc catggggaga gatgctagga gtggggggcg accggaacga ggagaaaagt 420
aaagagagca acggggctag cgagtggatg ttccgcctcc ccggggagcc gaggtaggct 480
tatcccggag tgccgggcaa gtcccccat gcccgttcca cggctccctt ccagaagggc 540
gggggtggcc tggagcgtcg ggccccagca gtccgtggaa ttcaccctct taccgcttct 600
ccacactcct tccccctgc gggctcccc cataagatga cgaggacccc actcatcggg 660
gtccgcagtc tcatccttct ttacctgat gccggcatgg tccagcctc ctgctggcg 720
ccggtcgggc aacattccga ggggaccgtc cctcggyaat ggcgaatggg acccagaact 780
ctctcagaat ctagggagat ctcccagatt cgagagaaaa ctggctctcc cttagccatc 840
cgagtggacg tctgtcctcc tacggatgcc caggtcggac cgcggggagg tggagatgcc 900
atgccgaccc gaagaggaaa gaaggactcg gacgcgaacc cgtgagtgga aacctcgttc 960
tttattgggg agtacactcg aggagtggaa ggcggggagg gcggggagga gtgttaccy 1020
acgggaactg ctgagtgcc ctgacgtcca gaccctcccc cgtccgggag aatggagatt 1080
cgggaacgtg aagcatggtg gggacgaagc ccctccggg cgtccctc ggatctccgc 1140
cgggtgggtt cacatccca acccggggc cggctgttct tcttttctt tgctcgtctt 1200
ccccggtcag cctcccaggt tctcttctt ctccctgct gaggttctt cctccggcgg 1260
ccagttgctt cttcttgttc tcgagggect tccttcgtcg gtgatccgc ctctcctcgt 1320
cggtgaaacc cgacttgtgc ggttttcccc taggtccgga atcgacctcc atttgatccg 1380
tccgggcctt ctctgttggg ggtgctccct ccccgctcct tcccttctt atgattccga 1440
tgatattccc caaccagga ttgtcatcct cgagtttctt gaggccctt tgggtcttcc 1500
ggagcttcc ttcgagatcc tctcgatttt ttcttagtt caccacttc tcgaggatct 1560
cttccctccc acctcgttt ctcttctgct cggcggggccc catctcgact gggggggcggc 1620
gtcctcagta ctctcttact ttctcaagaa aagaggagac tgctggtcgc ccgccccctg 1680
gttcgag 1687

```

5

<210> 22

<211> 642

<212> ADN

10 <213> virus de la hepatitis D

<220>

<221> CDS

<222> (1)..(692)

15 <223> ADNc de LHD

<400> 22

atg ggc ccc gcc gag cag aag aga aaa cga ggt ggg agg gaa gag atc	48
Met Gly Pro Ala Glu Gln Lys Arg Lys Arg Gly Gly Arg Glu Glu Ile	
1 5 10 15	
ctc gag aag tgg gtg gaa cta aga aaa aat cga gag gat ctc gaa agg	96
Leu Glu Lys Trp Val Glu Leu Arg Lys Asn Arg Glu Asp Leu Glu Arg	
20 25 30	
aag ctc cgg aag acc cag aag ggc ctc aag aaa ctc gag gat gac aat	144
Lys Leu Arg Lys Thr Gln Lys Gly Leu Lys Lys Leu Glu Asp Asp Asn	
35 40 45	
ccc tgg ttg ggg aat atc atc gga atc ata aga aag gga aag gac ggg	192
Pro Trp Leu Gly Asn Ile Ile Gly Ile Ile Arg Lys Gly Lys Asp Gly	
50 55 60	
gag gga gca ccc cca acg aag aag gcc cgg acg gat caa atg gag gtc	240
Glu Gly Ala Pro Pro Thr Lys Lys Ala Arg Thr Asp Gln Met Glu Val	
65 70 75 80	
gat tcc gga cct agg gga aag ccg cac aag tcg ggt ttc acc gac gag	288
Asp Ser Gly Pro Arg Gly Lys Pro His Lys Ser Gly Phe Thr Asp Glu	
85 90 95	
gag agg cgg gat cac cga cga agg aag gcc ctc gag aac aag aag aag	336
Glu Arg Arg Asp His Arg Arg Arg Lys Ala Leu Glu Asn Lys Lys Lys	
100 105 110	
caa ctg gcc gcc gga ggg aag aac ctc agc agg gaa gaa gaa gag gaa	384
Gln Leu Ala Ala Gly Gly Lys Asn Leu Ser Arg Glu Glu Glu Glu Glu	
115 120 125	
ctc ggg agg ctg acc ggg gaa gac gag caa agg aaa aga aga aca gcc	432
Leu Gly Arg Leu Thr Gly Glu Asp Glu Gln Arg Lys Arg Arg Thr Ala	
130 135 140	
ggc ccg cgg gtt ggg gat gtg aac cca ccc ggc gga gat ccg agg gga	480
Gly Pro Arg Val Gly Asp Val Asn Pro Pro Gly Gly Asp Pro Arg Gly	
145 150 155 160	
gcg ccc gga ggg ggc ttc gtc ccc acc atg ctt cac gtt ccc gaa tct	528
Ala Pro Gly Gly Gly Phe Val Pro Thr Met Leu His Val Pro Glu Ser	
165 170 175	
cca ttc tcc cgg acg ggg gag ggt ctg gac gtc agg ggc act cag cag	576
Pro Phe Ser Arg Thr Gly Glu Gly Leu Asp Val Arg Gly Thr Gln Gln	
180 185 190	
ttc ccg tgg ggt aac act cct ccc cgc cct ccc cgc ctt cca ctc ctc	624
Phe Pro Trp Gly Asn Thr Pro Pro Arg Pro Pro Arg Leu Pro Leu Leu	
195 200 205	
gag tgt act ccc caa taa	642
Glu Cys Thr Pro Gln	
210	

<210> 23

<211> 213

5 <212> PRT

<213> virus de la hepatitis D

<400> 23

Met Gly Pro Ala Glu Gln Lys Arg Lys Arg Gly Gly Arg Glu Glu Ile
 1 5 10 15
 Leu Glu Lys Trp Val Glu Leu Arg Lys Asn Arg Glu Asp Leu Glu Arg
 20 25 30
 Lys Leu Arg Lys Thr Gln Lys Gly Leu Lys Lys Leu Glu Asp Asp Asn
 35 40 45
 Pro Trp Leu Gly Asn Ile Ile Gly Ile Ile Arg Lys Gly Lys Asp Gly
 50 55 60
 Glu Gly Ala Pro Pro Thr Lys Lys Ala Arg Thr Asp Gln Met Glu Val
 65 70 75 80
 Asp Ser Gly Pro Arg Gly Lys Pro His Lys Ser Gly Phe Thr Asp Glu
 85 90 95
 Glu Arg Arg Asp His Arg Arg Arg Lys Ala Leu Glu Asn Lys Lys Lys
 100 105 110
 Gln Leu Ala Ala Gly Gly Lys Asn Leu Ser Arg Glu Glu Glu Glu Glu
 115 120 125
 Leu Gly Arg Leu Thr Gly Glu Asp Glu Gln Arg Lys Arg Arg Thr Ala
 130 135 140
 Gly Pro Arg Val Gly Asp Val Asn Pro Pro Gly Gly Asp Pro Arg Gly
 145 150 155 160
 Ala Pro Gly Gly Gly Phe Val Pro Thr Met Leu His Val Pro Glu Ser
 165 170 175
 Pro Phe Ser Arg Thr Gly Glu Gly Leu Asp Val Arg Gly Thr Gln Gln
 180 185 190
 Phe Pro Trp Gly Asn Thr Pro Pro Arg Pro Pro Arg Leu Pro Leu Leu
 195 200 205
 Glu Cys Thr Pro Gln
 210

5

<210> 24

<211> 585

<212> ADN

<213> virus de la hepatitis D

10

<220>

<221> CDS

<222> (1)..(585)

<223> ADNc de sHD

<400> 24

atg ggc ccc gcc gag cag aag aga aaa cga ggt ggg agg gaa gag atc	48
Met Gly Pro Ala Glu Gln Lys Arg Lys Arg Gly Gly Arg Glu Glu Ile	
1 5 10 15	
ctc gag aag tgg gtg gaa cta aga aaa aat cga gag gat ctc gaa agg	96
Leu Glu Lys Trp Val Glu Leu Arg Lys Asn Arg Glu Asp Leu Glu Arg	
20 25 30	
aag ctc cgg aag acc cag aag ggc ctc aag aaa ctc gag gat gac aat	144
Lys Leu Arg Lys Thr Gln Lys Gly Leu Lys Lys Leu Glu Asp Asp Asn	
35 40 45	
ccc tgg ttg ggg aat atc atc gga atc ata aga aag gga aag gac ggg	192
Pro Trp Leu Gly Asn Ile Ile Gly Ile Ile Arg Lys Gly Lys Asp Gly	
50 55 60	
gag gga gca ccc cca acg aag aag gcc cgg acg gat caa atg gag gtc	240
Glu Gly Ala Pro Pro Thr Lys Lys Ala Arg Thr Asp Gln Met Glu Val	
65 70 75 80	
gat tcc gga cct agg gga aag ccg cac aag tcg ggt ttc acc gac gag	288
Asp Ser Gly Pro Arg Gly Lys Pro His Lys Ser Gly Phe Thr Asp Glu	
85 90 95	
gag agg cgg gat cac cga cga agg aag gcc ctc gag aac aag aag aag	336
Glu Arg Arg Asp His Arg Arg Arg Lys Ala Leu Glu Asn Lys Lys Lys	
100 105 110	
caa ctg gcc gcc gga ggg aag aac ctc agc agg gaa gaa gaa gag gaa	384
Gln Leu Ala Ala Gly Gly Lys Asn Leu Ser Arg Glu Glu Glu Glu Glu	
115 120 125	
ctc ggg agg ctg acc ggg gaa gac gag caa agg aaa aga aga aca gcc	432
Leu Gly Arg Leu Thr Gly Glu Asp Glu Gln Arg Lys Arg Arg Thr Ala	
130 135 140	
ggc ccg cgg gtt ggg gat gtg aac cca ccc ggc gga gat ccg agg gga	480
Gly Pro Arg Val Gly Asp Val Asn Pro Pro Gly Gly Asp Pro Arg Gly	
145 150 155 160	
gcg ccc gga ggg ggc ttc gtc ccc acc atg ctt cac gtt ccc gaa tct	528
Ala Pro Gly Gly Gly Phe Val Pro Thr Met Leu His Val Pro Glu Ser	
165 170 175	
cca ttc tcc cgg acg ggg gag ggt ctg gac gtc agg ggc act cag cag	576
Pro Phe Ser Arg Thr Gly Glu Gly Leu Asp Val Arg Gly Thr Gln Gln	
180 185 190	
ttc ccg tag	585
Phe Pro	
195	

5

<210> 25

<211> 194

<212> PRT

<213> virus de la hepatitis D

<400> 25

```

Met Gly Pro Ala Glu Gln Lys Arg Lys Arg Gly Gly Arg Glu Glu Ile
 1          5          10          15
Leu Glu Lys Trp Val Glu Leu Arg Lys Asn Arg Glu Asp Leu Glu Arg
      20          25          30
Lys Leu Arg Lys Thr Gln Lys Gly Leu Lys Lys Leu Glu Asp Asp Asn
      35          40          45
Pro Trp Leu Gly Asn Ile Ile Gly Ile Ile Arg Lys Gly Lys Asp Gly
      50          55          60
Glu Gly Ala Pro Pro Thr Lys Lys Ala Arg Thr Asp Gln Met Glu Val
      65          70          75          80
Asp Ser Gly Pro Arg Gly Lys Pro His Lys Ser Gly Phe Thr Asp Glu
      85          90          95
Glu Arg Arg Asp His Arg Arg Arg Lys Ala Leu Glu Asn Lys Lys Lys
      100          105          110
Gln Leu Ala Ala Gly Gly Lys Asn Leu Ser Arg Glu Glu Glu Glu Glu
      115          120          125
Leu Gly Arg Leu Thr Gly Glu Asp Glu Gln Arg Lys Arg Arg Thr Ala
      130          135          140
Gly Pro Arg Val Gly Asp Val Asn Pro Pro Gly Gly Asp Pro Arg Gly
      145          150          155          160
Ala Pro Gly Gly Gly Phe Val Pro Thr Met Leu His Val Pro Glu Ser
      165          170          175
Pro Phe Ser Arg Thr Gly Glu Gly Leu Asp Val Arg Gly Thr Gln Gln
      180          185          190

Phe Pro

```

5

<210> 26

<211> 1680

<212> ADN

10 <213> virus de la hepatitis D

<220>

<221> cepa

<222> (1) .. (1680)

15 <223> secuencia genómica completa de dFr644

<400> 26

```

atgggccaaag agccgacgaa gaggtcgaag gaggagaagg agtgagtccc caaggggttc 60
gaagccaaaag agtggaaca ctcgaggaag agggagtccc aagagggcga ggggagcttt 120
caggagggag ggtgaatccc cgagrcgctg gagactcccc gaatgaggaa agtggggata 180

ggctaggtag aaaagagcga gcctcccaaa tcgagattgc cctgaacctt tcragtttgg 240
agtcattccga gcttaggggtt gaattgtagcg cagagggagg aagccaccag gagccggaga 300
caagacacct ccagaggacc ccttcagcga acagaagagt ctgtccccag atgagaagag 360
cgtagctaga gggggagatg ctgggagtag ggggagatcg aagcgaggag gaaagtaaag 420
atggcaacgg ggctagcagg aggtgtgtcc gccccccggc ggggccgagt gaggcttatt 480
ccggggaact cggactggtc ccggactgcc ggctccaggg gccactcca aaggaccgag 540
ggcaggactt ggagcaccgg gaattccgag caactccatg gtagactccg tcccccttct 600
ccacactcct tccccccggc ggcccccccg taaagtggag aacccactc cgagggttcc 660
gcgccctcct ccctttctta cctgtggccg gcatggtccc agcctcctcg cggcgccggc 720
cgggcaacat tccgagggga ccgtcccccg gtaatggcag atgggacca ggcttcccc 780
ggattccctc atggggatcg agggggagag ctggcactcc cttagccatc cgagtggacg 840
tctgtcctcc ttccgatgcc caggtcggac caccgggagg tggagatgoc atgccacce 900
gaagagggaa gaaggagacc ggacgcaaac ctgtgagtgg aactctcdtc ctttattggg 960
gggtacactc gaggagtgga aggcggggag gggggggcgg tctttgtccc tatggaaatt 1020
gctggtctcc cctgatgtcc agtccatccc cgtgtctggt gaatggagac tccggaacac 1080
ctagcatcct agggacaaat ccgcccccg ggcctcccct cggacttcct ccgggagggt 1140
tcacaccccc atcctgctgg ccggccgctc ttctttctct tttctcgtct tcaacgatca 1200
acctcctgag ttccccctct tcctcctcgc tgagcttctt ccctccggag gacagttgct 1260
tcttcttgtt ctcgagggcc ttctttcttc ggtgatccct cctctcctcg tcggtgaatc 1320
ctcccttgcg actcttcttc ccgggaccgg agtcgacctc catctgatct gtccgggtc 1380
tcttcgctgg gggagctccc tcccctgctt tcccccttct tattattccg aggatgttcc 1440
ccagccaggg attttcatcc tcgagtctct tgatgttctt cctcgtcttc cggatcctct 1500
tttcgaggtc ttccagatct tttcttgctt cgaccactt tgagaggaca tcctcccgtc 1560
ctccccctcg ttctctcttc gcatcggact ggctcatctc ggcgagggcg gcgatcctca 1620
gtgctcttac tctttacagt agaaagagga gactgctgga tgccccgccc gggcycgagc 1680

```

5

<210> 27

<211> 645

<212> ADN

<213> virus de la hepatitis D

10

<220>

<221> CDS

<222> (1)..(645)

<223> ADNc de LHD

15

<400> 27

atg agc cag tcc gat gcg aag aga gaa cga agg gga gga cgg gag gat	48
Met Ser Gln Ser Asp Ala Lys Arg Glu Arg Arg Gly Gly Arg Glu Asp	
1 5 10 15	
gtc ctc tca aag tgg gtc gaa gca aga aaa gat ctg gaa gac ctc gaa	96
Val Leu Ser Lys Trp Val Glu Ala Arg Lys Asp Leu Glu Asp Leu Glu	
20 25 30	
aag agg atc cgg aag acg agg agg aac atc aag aga ctc gag gat gaa	144
Lys Arg Ile Arg Lys Thr Arg Arg Asn Ile Lys Arg Leu Glu Asp Glu	
35 40 45	
aat ccc tgg ctg ggg aac atc ctc gga ata ata aga aag ggg aag gca	192
Asn Pro Trp Leu Gly Asn Ile Leu Gly Ile Ile Arg Lys Gly Lys Ala	
50 55 60	
ggg gag gga gct ccc cca gcg aag aga gcc cga aca gat cag atg gag	240
Gly Glu Gly Ala Pro Pro Ala Lys Arg Ala Arg Thr Asp Gln Met Glu	
65 70 75 80	
gtc gac tcc ggt ccg ggg aag aag agt cgc aag gga gga ttc acc gac	288
Val Asp Ser Gly Pro Gly Lys Lys Ser Arg Lys Gly Gly Phe Thr Asp	
85 90 95	
gag gag agg agg gat cac cga aga agg aag gcc ctc gag aac aag aag	336
Glu Glu Arg Arg Asp His Arg Arg Arg Lys Ala Leu Glu Asn Lys Lys	
100 105 110	
aag caa ctg tcc tcc gga ggg aag aag ctc agc gag gag gaa gaa ggg	384
Lys Gln Leu Ser Ser Gly Gly Lys Lys Leu Ser Glu Glu Glu Glu Gly	
115 120 125	
gaa ctc agg agg ttg atc gtt gaa gac gag aaa aga gaa aga aga gcg	432
Glu Leu Arg Arg Leu Ile Val Glu Asp Glu Lys Arg Glu Arg Arg Ala	
130 135 140	
gcc gcc ccg cag gat ggg ggt gtg aac cct ccc gga gga agt ccg agg	480
Ala Gly Pro Gln Asp Gly Gly Val Asn Pro Pro Gly Gly Ser Pro Arg	
145 150 155 160	
gga gcg ccc ggg gcc gga ttt gtc cct agg atg cta ggt gtt ccg gag	528
Gly Ala Pro Gly Gly Gly Phe Val Pro Arg Met Leu Gly Val Pro Glu	
165 170 175	
tct cca ttc acc aga cac ggg gat gga ctg gac atc agg gga gac cag	576
Ser Pro Phe Thr Arg His Gly Asp Gly Leu Asp Ile Arg Gly Asp Gln	
180 185 190	
caa ttt cca tgg gga caa aga ccg ccc ccc cct ccc cgc ctt cca ctc	624
Gln Phe Pro Trp Gly Gln Arg Pro Pro Pro Pro Arg Leu Pro Leu	
195 200 205	
ctc gag tgt acc ccc caa taa	645
Leu Glu Cys Thr Pro Gln	
210 215	

<211> 214

<212> PRT

<213> virus de la hepatitis D

5 <400> 28

```

Met Ser Gln Ser Asp Ala Lys Arg Glu Arg Arg Gly Gly Arg Glu As
 1           5           10           15

Val Leu Ser Lys Trp Val Glu Ala Arg Lys Asp Leu Glu Asp Leu Gl
      20           25           30

Lys Arg Ile Arg Lys Thr Arg Arg Asn Ile Lys Arg Leu Glu Asp Gl
      35           40           45

Asn Pro Trp Leu Gly Asn Ile Leu Gly Ile Ile Arg Lys Gly Lys Al
 50           55           60

Gly Glu Gly Ala Pro Pro Ala Lys Arg Ala Arg Thr Asp Gln Met Gl
 65           70           75           8

```

```

Val Asp Ser Gly Pro Gly Lys Lys Ser Arg Lys Gly Gly Phe Thr Asp
      85           90           95

Glu Glu Arg Arg Asp His Arg Arg Arg Lys Ala Leu Glu Asn Lys Lys
     100           105           110

Lys Gln Leu Ser Ser Gly Gly Lys Lys Leu Ser Glu Glu Glu Glu Gly
     115           120           125

Glu Leu Arg Arg Leu Ile Val Glu Asp Glu Lys Arg Glu Arg Arg Ala
     130           135           140

Ala Gly Pro Gln Asp Gly Gly Val Asn Pro Pro Gly Gly Ser Pro Arg
    145           150           155           160

Gly Ala Pro Gly Gly Gly Phe Val Pro Arg Met Leu Gly Val Pro Glu
     165           170           175

Ser Pro Phe Thr Arg His Gly Asp Gly Leu Asp Ile Arg Gly Asp Gln
     180           185           190

Gln Phe Pro Trp Gly Gln Arg Pro Pro Pro Pro Pro Arg Leu Pro Leu
     195           200           205

Leu Glu Cys Thr Pro Gln
     210

```

<210> 29

10 <211> 588

<212> ADN

<213> virus de la hepatitis D

<220>

<221> CDS

<222> (1)..(588)

5 <223> ADNc de sHD

<400> 29

atg agc cag tcc gat gcg aag aga gaa cga agg gga gga cgg gag gat	48
Met Ser Gln Ser Asp Ala Lys Arg Glu Arg Arg Gly Gly Arg Glu Asp	
1 5 10 15	
gtc ctc tca aag tgg gtc gaa gca aga aaa gat ctg gaa gac ctc gaa	96
Val Leu Ser Lys Trp Val Glu Ala Arg Lys Asp Leu Glu Asp Leu Glu	
20 25 30	
aag agg atc cgg aag acg agg agg aac atc aag aga ctc gag gat gaa	144
Lys Arg Ile Arg Lys Thr Arg Arg Asn Ile Lys Arg Leu Glu Asp Glu	
35 40 45	
aat ccc tgg ctg ggg aac atc ctc gga ata ata aga aag ggg aag gca	192
Asn Pro Trp Leu Gly Asn Ile Leu Gly Ile Ile Arg Lys Gly Lys Ala	
50 55 60	
ggg gag gga gct ccc cca gcg aag aga gcc cga aca gat cag atg gag	240
Gly Glu Gly Ala Pro Pro Ala Lys Arg Ala Arg Thr Asp Gln Met Glu	
65 70 75 80	
gtc gac tcc ggt ccg ggg aag aag agt cgc aag gga gga ttc acc gac	288
Val Asp Ser Gly Pro Gly Lys Lys Ser Arg Lys Gly Gly Phe Thr Asp	
85 90 95	
gag gag agg agg gat cac cga aga agg aag gcc ctc gag aac aag aag	336
Glu Glu Arg Arg Asp His Arg Arg Arg Lys Ala Leu Glu Asn Lys Lys	
100 105 110	
aag caa ctg tcc tcc gga ggg aag aag ctc agc gag gag gaa gaa ggg	384
Lys Gln Leu Ser Ser Gly Gly Lys Lys Leu Ser Glu Glu Glu Glu Gly	
115 120 125	
gaa ctc agg agg ttg atc gtt gaa gac gag aaa aga gaa aga aga gcg	432
Glu Leu Arg Arg Leu Ile Val Glu Asp Glu Lys Arg Glu Arg Arg Ala	
130 135 140	
gcc ggc ccg cag gat ggg ggt gtg aac cct ccc gga gga agt ccg agg	480
Ala Gly Pro Gln Asp Gly Gly Val Asn Pro Pro Gly Gly Ser Pro Arg	
145 150 155 160	
gga gcg ccc ggg ggc gga ttt gtc cct agg atg cta ggt gtt ccg gag	528
Gly Ala Pro Gly Gly Gly Phe Val Pro Arg Met Leu Gly Val Pro Glu	
165 170 175	
tct cca ttc acc aga cac ggg gat gga ctg gac atc agg gga gac cag	576
Ser Pro Phe Thr Arg His Gly Asp Gly Leu Asp Ile Arg Gly Asp Gln	
180 185 190	
caa ttt cca tag	588
Gln Phe Pro	
195	

<210> 30

<211> 195

5 <212> PRT

<213> virus de la hepatitis D

<400> 30

```

Met Ser Gln Ser Asp Ala Lys Arg Glu Arg Arg Gly Gly Arg Glu Asp
 1          5          10          15

Val Leu Ser Lys Trp Val Glu Ala Arg Lys Asp Leu Glu Asp Leu Glu
      20          25          30

Lys Arg Ile Arg Lys Thr Arg Arg Asn Ile Lys Arg Leu Glu Asp Glu
      35          40          45

Asn Pro Trp Leu Gly Asn Ile Leu Gly Ile Ile Arg Lys Gly Lys Ala
      50          55          60

Gly Glu Gly Ala Pro Pro Ala Lys Arg Ala Arg Thr Asp Gln Met Glu
      65          70          75          80

Val Asp Ser Gly Pro Gly Lys Lys Ser Arg Lys Gly Gly Phe Thr Asp
      85          90          95

Glu Glu Arg Arg Asp His Arg Arg Arg Lys Ala Leu Glu Asn Lys Lys
      100          105          110

Lys Gln Leu Ser Ser Gly Gly Lys Lys Leu Ser Glu Glu Glu Glu Gly
      115          120          125

Glu Leu Arg Arg Leu Ile Val Glu Asp Glu Lys Arg Glu Arg Arg Ala
      130          135          140

Ala Gly Pro Gln Asp Gly Gly Val Asn Pro Pro Gly Gly Ser Pro Arg
      145          150          155          160

Gly Ala Pro Gly Gly Gly Phe Val Pro Arg Met Leu Gly Val Pro Glu
      165          170          175

Ser Pro Phe Thr Arg His Gly Asp Gly Leu Asp Ile Arg Gly Asp Gln
      180          185          190

Gln Phe Pro
      195

```

5

<210> 31

<211> 25

<212> ADN

<213> virus de la hepatitis D

10

<400> 31

gaggaaagaa ggacgcgaga cgcaa

<210> 32

<211> 21

<212> ADN

<213> virus de la hepatitis D

5

<400> 32

accccctcga aggtggatcg a 21

<210> 33

10 <211> 22

<212> ADN

<213> virus de la hepatitis D

<400> 33

15 catgccgacc cgaagaggaa ag 22

<210> 34

<211> 24

<212> ADN

20 <213> virus de la hepatitis D

<400> 34

gaaggaaggc cctcgagaac aaga 24

25 <210> 35

<211> 24

<212> ADN

<213> virus de la hepatitis D

30 <400> 35

ctccagagga ccccttcagc gaac 24

<210> 36

<211> 24

35 <212> ADN

<213> virus de la hepatitis D

<400> 36

cccgcgggtt ggggatgtga accc 24

<210> 37

<211> 23 .

5 <212> ADN

<213> virus de la hepatitis D

<400> 37

gtacactcga ggagtggaag gcg 23

10

<210> 38

<211> 21

<212> ADN

<213> virus de la hepatitis D

15

<400> 38

tctgttcgct gaaggggtcc t 21

<210> 39 <211> 22 <212> ADN

20 <213> virus de la hepatitis D

<400> 39

ccagaggacc ccttcagcga ac 22

25 <210> 40

<211> 20

<212> ADN

<213> virus de la hepatitis D

30 <400> 40

aacaccctcc tgctagccccc 20

<210> 41

<211> 22

35 <212> ADN

<213> virus de la hepatitis D

<400> 41

ccggaggtcc tctctctct cc 22

<210> 42

<211> 23

5 <212> ADN

<213> virus de la hepatitis D

<400> 42

gttcgcgtcc gagtccttct ttc 23

10

<210> 43

<211> 20

<212> ADN

<213> virus de la hepatitis D

15

<400> 43

gagctttctt cgattcggac 20

<210> 44

20 <211> 20

<212> ADN

<213> virus de la hepatitis D

<400> 44

25 gactggtccc ctcattgtcc 20

<210> 45

<211> 22

<212> ADN

30 <213> virus de la hepatitis D

<400> 45

gttcgctgaa ggggtctctt gg 22

35 <210> 46

<211> 24

<212> ADN

<213> virus de la hepatitis D

<400> 46

tcacccctga gtctcttgat ggctc 24

5 <210> 47

<211> 20

<212> ADN

<213> virus de la hepatitis D

10 <400> 47

aacatccact cgttagccccc 20

<210> 48

<211> 396

15 <212> ADN

<213> virus de la hepatitis D

<400> 48

```

catgccgacc cgaagaggaa agaaggattc ggacgcaaac ctgtgagtgg aagttcgctc 60
tttattgggg ggtacactcg aggagtggaa ggcggggagg gggggtcgga ctaggtccct 120
atggaaactg tcggtctcct cggatgtcga gtccctctcc cgttctggag aagggggact 180
ccgggacacc tagcagttga ggaacgaagc cgcgcccggg cgctcccctc ggtgggccct 240
cgggaggggt cacatcccca acccgcgggc cggetattct tcttgcctct tgcctgcctc 300
cggcggtcaa cttcctgagt tcctctctct cctccttgct gagggacttt cctcctgcgg 360
acagctgctt cttcttggtc tcgagggcct tccttc 396

```

20

<210> 49

<211> 397

<212> ADN

25 <213> virus de la hepatitis D

<400> 49

```

catgccgacc cgaagaggaa agaaggactc ggacgcgaac tcgtgagtgg aaacttattc 60
ctttactggg gagtacactc gaggagtggg aggcggggag gccgggggtc cgggytacc 120
cacgggtact gctggttccc cctcacgtcc agtccctccc ccgtcctgga gaaggagac 180
tcrggaacac ctacatcttg agggacaaag ccgcccccg gcgtcccctc cggagatgtt 240
ccgggtgggt tcacatcccc aaccgcgggg ccggctactc ttctttctct tctctcgtct 300
tcctcgggtc acttcggag ttctctctcc tcctccttgc tragtgactt tcctcccgcg 360
ctcagttggt tcttcttggt ctcgagggcc tccttc 397

```

30

<210> 50

<211> 397

<212> ADN

<213> virus de la hepatitis D

5

<400> 50

```

catgccgacc cgaagaggaa agaaggactc ggacgcgaac ccgtgagtgg aaacctcggt 60
ctttattggg gagtacactc gaggagtggg aggcggggag ggcggggagg agtgttaccc 120
yacgggaact gctgagtgcc cctgacgtcc agaccctccc ccgtccggga gaatggagat 180
tcgggaacgt gaagcatggt ggggacgaag cccctccggg gcgtccccct cggatctccg 240
ccgggtgggt tcacatcccc aaccgcgggg ccggtgttcc ttcttttctc ttgctcgtct 300
tcccgcgtca gctcccgag ttctcttctc tcttccctgc tgaggttctt ccctccggcg 360
gccagttgct tcttcttggt ctcgagggcc ttccttc

```

10 <210> 51

<211> 410

<212> ADN

<213> virus de la hepatitis D

15 <400> 51

```

aygccatgcc gaccgaaga aggaagaag gactcggacg cgaaccctg agtggaaactc 60
tttatctttt actggggagt acactcgagg agtgggaagg ggggaggccg gggccccggg 120

```

```

tctaccacg ggaattgctg gtttccctt acgtccagtc cacttcccgt cctggagaag 180
ggagactcgg gaacatctag catttgaggg acaaagccgc ccccgggcgc tcccctcgga 240
ggccccctcg gtgggttca acccccacc ccggggccgg ctactcttct tccccttctc 300
tcgtcttctc cgtcaactt ccggagttcc tcttctctc ccctgctgag tgactttccc 360
cccgcgtca gctgattctt cttgttctcg agggctttcc ttcttcggtg 410

```

<210> 52

20 <211> 397

<212> ADN

<213> virus de la hepatitis D

<400> 52

25


```

catgccgacc cgaagaggaa agaaggactc ggacgcgaac ccgtgagtgg aaccttattc 60
ctttattggg gagtacactc gaggagtgga aggcggggag acggggctcc cgggcttacc 120
yacgggaact gctggttccc ccttatgtcc agtccctctc ccgtcctggt gaaggragac 180
tcgggaacgt tcagcatttg agggacaaag ccgcccccg ggcctcccct cggaggtccc 240
tgtggtgggt tcacatcccc aaccgcggg ccggtactc ttctttccct tctctcgtct 300
tcytcgggtc acctcccagg ttctctttct tctccttgcc tgagtgaact tccycccgcg 360
ttcagctggt tcttcttggt ctcgagggcc ttccttc 397

```

<210> 53

<211> 399

5 <212> ADN

<213> virus de la hepatitis D

<400> 53

```

catgccgacc cgaagaagga aagaaggact cggacgcgaa cccgtgagtg gaactcttta 60
tcttttactg gggagtacac tcgaggagtg gaaggcgggg aggcgggggt cccgggtcta 120
cccacgggaa ttgctgggtt ccccttacgt ccagtccact tcccgctctg gagaaggag 180
actcgggaac atctagcatt tgaggggacaa agccgcccc ggcgctccc ctcggaggtc 240
cctcgggtgg gttcacatcc ccaacccgcg ggccggtac tcttctttcc cttctctcgt 300
cttcctcggg caacttccgg agttcctctt cctcctccct gctgagtgaac tttccccccg 360
cgctcagctg attcttcttg ttctcgaggg ctttccttc 399

```

10

<210> 54

<211> 397

<212> ADN

15 <213> virus de la hepatitis D

<400> 54

```

catgccgacc cgaagaggaa agaaggattc ggacgcaaac ctgtgagtgg aaattctctt 60
ctttattggg gagtacactc gaggagtgga aggcggggag ggggggtcgg actaggaacc 120
tatgggaact gctggtttcc tcggatgtcc agtccatccc ccgtcctgga gaatggtgac 180
tccgggactc ctagcagttg aggaacgaag ccgcccccg ggcctcccct cgaaggcgcg 240
gggggagggg tcacatcccc aaccgcggg ccggtgttcc ttcttgctct tttctcgtcc 300
tcgacgggtc gcctcgcgag ttctctttct tcttcttgcc cgagggttct tctcccgcg 360
gagagttgct tcctcttggt ctcgagggcc ttccttc 397

```

20

<210> 55

<211> 397

<212> ADN

<213> virus de la hepatitis D

25

<400> 55

```

catgccgacc cgaagaggaa agaaggactc ggacgcgaac tcgtgagtgg aaccttaatc 60
ctttattggg gagtacactc gaggagtggg aggcggggag gccgggggtcc cgagcttacc 120
tacgggaact gctggtttcc tcttatgtcc agtccctctc ccgtccctaga gaaggagac 180
tcgggaacgc ccagcatttg agggacaaag ccgcccccg ggcctcccct cggagggtcc 240
tcgggwggtt tcacatcccc aaccgcggg ccggtactc ttctttccct tctctcgtct 300
tcctcgggtca acttcggag ttctcttctc tcctccctgc tgaggttctt tcctcccgcg 360
gagagttgct tcctcttgtt ctcgagggcc ttccttc 397

```

<210> 56

5 <211> 397

<212> ADN

<213> virus de la hepatitis D

<400> 56

10

```

catgccgacc cgaagaggaa agaaggactc ggacgcgaac tcgtgagtgg acacctcgtt 60
ctttattggg gagtacactc gaggagtggg aggcggggag gccggggagg agtggttacc 120
tatggaaact gctggttwcc cctcacgtcc agaccttctc ccgtccggga gaatggagat 180
tcgggaacga gatgcatggt tgggacgaag cccctcccg ggcctcccct cggacttccc 240
tcgggtgggt tcacatcccc aaccgcggg ccggtattc ttctttccct ttgctcgtca 300
tcctcgggtca gcctcaggag ttctcttctc tcttcttgc tgaggttctt tcctcccgcg 360
gagagttgct tcctcttgtt ctcgagggcc ttccttc 397

```

<210> 57

<211> 396

15 <212> ADN

<213> virus de la hepatitis D

<400> 57

```

catgccgacc cgaagaggaa agaaggattc ggacgcaaac ctgtgagtgg aagctttctc 60
tttattgggg ggtacactcg aggagtggaa ggcggggagg ggggggtcga ctaggtcccy 120
atggaaactg ccggtttcct cggacgtcga gtccctctcc cgttctggag aagggggact 180
ccgggactcc tagcagttga ggaacgaagc cgcccccg ggcctcccct ggtggtccct 240
cgggagggtt cacatcccca acctgcgggc cggctactct tcttgccttt tgctcgtctt 300
cgacggtcaa cctctcgagt tcctcttctc cctccttgc gaggttcttt cctcccgcg 360
agagttgctt cctcttgttc tcgagggcct ttccttc 396

```

20

<210> 58

<211> 398

<212> ADN

25 <213> virus de la hepatitis D

<400> 58

```

catgccgacc cgaagaggaa agaaggagac cggacgcaaa cctgtgagtg gaactctcdt 60
cctttattgg ggggtacact cgaggagtgg aaggcgggga ggggggggcg gtctttgtcc 120
ctatggaaat tgctgggtctc ccctgatgtc cagtccatcc ccgtgtctgg tgaatggaga 180
ctccggaaca cctagcatcc tagggacaaa tccgcccccg ggcgctcccc tcggacttcc 240
tccgggaggg ttcacacccc catcctgcgg gccggccgct cttctttctc tttctcgtc 300
ttcaacgata aacctcctga gttccccttc ttcctcctcg ctgagcttct tccctccgga 360
ggacagttgc ttcttcttgt tctcgagggc cttccttc 398

```

5 <210> 59

<211> 10

<212> PRT

<213> virus de la hepatitis D

10 <400> 59

```

Arg Leu Pro Leu Leu Glu Cys Thr Pro Gln
 1             5             10

```

<210> 60

15 <211> 9

<212> PRT

<213> virus de la hepatitis D

<400> 60

20

```

Trp Gly Pro Ser Pro Thr Pro Pro Pro
 1             5

```

<210> 61

<211> 9

25 <212> PRT

<213> virus de la hepatitis D

<400> 61

```

Trp Val Xaa Pro Gly Pro Arg Pro Pro
 1             5

```

30

<210> 62

<211> 9

<212> PRT

5 <213> virus de la hepatitis D

<400> 62

Trp	Val	Ser	Pro	Gly	Ala	Pro	Ser	Pro
1				5				

10

<210> 63

<211> 9

<212> PRT

<213> virus de la hepatitis D

15

<400> 63

Trp	Val	Asp	Pro	Gly	Pro	Arg	Pro	Pro
1				5				

20

<210> 64

<211> 9

<212> PRT

<213> virus de la hepatitis D

25

<400> 64

Trp	Gly	Asn	Thr	Pro	Pro	Arg	Pro	Pro
1				5				

<210> 65

30

<211> 9

<212> PRT

<213> virus de la hepatitis D

<400> 65

Trp Gly Gln Arg Pro Pro Pro Pro Pro
 1 5

5 <210> 66

<211> 1685

<212> ADN

<213> virus de la hepatitis D

10 <220>

<221> cepa

<222> (1)..(1685)

<223> secuencia genómica completa de Pt26

15 <400> 66

```

atggggccgca taccgagcga aggataccgg taagggggaga ggaggaatcc cgaggggatc 60
gcccaactaaa gagtagagga attctcggaa agctctccca agaagaacca gaatccctca 120
agagagagga gaatccccat gacgctggaa gagacttccg ggtaaccaag tcaaggagaa 180
ggaacggtag aaaagagcga gcctctcgat acgaaagggc cgcgagccta tcgagtttgg 240
agtcattccgg gccgaagggt tgagacgcac ccacagacgg gagcctccag gagggtttga 300
ggagaactca cctccagagg accccttcag cgaacagagg agtctctacc tcggaggaga 360
aagaccgtag cgatggggga gatgctagga gtaggcggcg accgaagcga ggaggaaagc 420
aaagaaagcc acgggggctag cgagtggatg ttccgcccc aaggggagccg agtgaggctt 480
atcccgggga actcggtaat tcgcccogaa atgaggactc cgggggtccc ttccgaagga 540
acggaggggg gtgactagga atcaggatcc agtggatccg tgggaattat cccgttccac 600
cccccgggca cactccttcc cccctgcggg ccccccata agatggcagg aaccactca 660
ttgggggtccg ctgttccatt ctttcttacc ttgtggccgg catgggtcca gcctcctcgc 720
tggcgccggc tgggcaacat tccgagggga ccgtccctcg gtaatggcga atgggaccca 780
gaactctctc tagattccca gagagaatcg agagaaaact ggctctccct tagccatccg 840
agtaggacgt ctgtcctcct acggatgcc aggtcggacc gcggggaggt ggagatgcca 900
tgccgaccgg aagaggaaag aagaacacgg acgcgaaccc gtaagtggaa ccctgatcct 960
ttattggggg gtacactcga ggagtggaaag gcgtgcccg gggggaaccg gatttaccta 1020
cgggaatccc tgggttcctc tgatgtccag tccctctccc gtccgagtga agggagactc 1080
cggaactcca gtcatttgag ggaagaaatc gccccgggc gctcccctcg gacttcctcc 1140
aggaggggtc acatccccaa cccgcggggc ggctactctt ctttcccttc tctcgtcttc 1200
ctcgggtcaac ctctgagtt cctcttcttc ctccctgctg aggttcttcc ctcccggga 1260
gagttgcttc ctctgttct cgagggcctt ccttctgcgg tgatcctccc tctccttgc 1320
ggtgaaccgg ctctgtgag gcctcttcat aggtccggag tcgacctcca tctgatctgt 1380
tcgggcccctc ttccggggg gagctccctc cccgtccttc ccttttctta tgattcccag 1440
gatgttcccc agccagggat tttcatctc gagtctcttg atggtcttcc tggccttccg 1500
gaggtctctc tcgagctcct ctgcctttct tcttgtggag atccactttt cgaggatttc 1560
ttcccttctt cccctccggc ccttctctga ctccgggttg ctcatcctcg gcgagggcga 1620
cggctctcag ttctctctat tctttcttga aagaggagac tgctggacca aacgcccag 1680
ccggg 1685

```

<210> 67

20 <211> 1688

<212> ADN

<213> virus de la hepatitis D

<220>

5 <221> cepa

<222> (1)..(1688)

<223> secuencia genómica completa de Pt62

<400> 67

10

```

atggggccgca agccgaacga aggattccgg taagggggaga ggaagaatcc cgaagggggtt 60
ccccactaaa gagtggaaga attctcggga agcttctccc aagaagaacc agaatcccc 120
aagagagaga atggatcccc atgacgtctg aagagactcc gggtaaccaa gtcaaggaga 180
aggaacggta gaaaagagcg agcctctcga tacgaaaggg ccgcgacctt tcaagtttgg 240
agtcattccg gccaaagggt tgaaaaatcc cacagacggg agccaccagg agggatctag 300
gagaatccac ctccagagga cccccctcaa tgaacagaag actctctacc tcggaggaaa 360
aagaccatag cgataggaag agatgctagg agtaggcggc gaccaaagcg aggaagaaag 420
taaagaaagc aacgggggcta gcgagtggat gttccgcccc aaggggagcc gagtgaggct 480
tatcccgggg aactcggcgt atcgtcccg aatgaggagc ccggatcccc ttccaaaaag 540
acggagaggg ggtgactagg aatcgggctc cgggtgatcc gtgggaccag cccgctccac 600
ctccgcggca cactccttcc cccctgcggg cccccccata agatggcagg aaccactca 660
ttgggggtcc ctgttccatt ctttcttacc ttgtggccgg catggtccca gcctcctcgc 720
tggcgcggc tgggcaacat tccgagggga ccgtccctcg gtaatggcga atgggacca 780
gaactctctc tagattccca gagagaatcg agagaaaact ggctctccct tagccatccg 840
agtaggacgt ctgtcctcct acggatgccc aggtcggacc gcgaggaggt ggagatgcca 900
tgccgacccg aagaggaaaag aagaacacgg acgcgaaccc gtaagtggaa ccctgatcct 960
ttattggggg gtacactcga ggagtgaag gcgctgcccg gggggagccg gattgacctt 1020
cgggaatccc cggtcgcctc tgatgtccag tccctcccc gtccgagaga agggagattc 1080
cggaactcca gtcatttgag ggacgaagcc gcccccgggc gctcccctcg gacttcctcc 1140
aggagggttc acatccccaa ccgcggggcc ggctactctt ctttgtcttt cgtcgtcttc 1200
aatggtcaac ctcttgagtt cctcttcttc ttcttctgtg aggtcttttc ccccgcgga 1260
gagttgggtt ttcttgttct ggagggcctt ccttctgcgg tggctctgcc tctccttgtc 1320
ggtgaacccg ctcttgtgag gtttcttctt aggtccggag tcgacctcca tctgatctgt 1380
tcgggccctc ttccgcgggg gagctccctc ccgctccttc cttttcttta tgattcccag 1440
gatgttcccc agccagggat tgtcatctc gagtctcttg atggtctttc tggccttccg 1500
gaggtctctc tcgagctctt ccgccttttt tcttgtggat acccactttt cgaggatata 1560
ttcccttctt cccctccggc ttttctctga ttccgattgg ctcatcctcg acgagggcga 1620
cggctcctcag ttctctctat tctttccttt tgaaagagga gactgctggt ccaaaccgcc 1680
gagtcggg 1688

```

REIVINDICACIONES

1. Moléculas de ácido nucleico aisladas, **caracterizadas por que** se seleccionan del grupo constituido por:

- 5 - el genoma complete de la variante del HDV denominada dFr45 que presenta la secuencia SEC ID N°: 1, y
- el genoma de una variante del HDV que presenta una divergencia genética $\leq 15\%$ con la secuencia SEC ID N°: 1.

10 2. Moléculas de ácido nucleico, **caracterizadas por que** comprenden al menos uno de los fragmentos de la secuencia de un HDV variante de acuerdo con la reivindicación 1, seleccionadas del grupo constituido por:

- a) los fragmentos R0 que presenta la secuencia SEC ID N°: 48
- b) el fragmento R1 que se extiende desde las posiciones 307 a 1289 del genoma del HDV,
- 15 c) el fragmento R2 que se extiende desde las posiciones 889 a 328 del genoma del HDV,
- d) el fragmento R3 que se extiende desde las posiciones 1486 a 452 del genoma del HDV,
- e) el fragmento R'1 que se extiende desde las posiciones 305 a 1161 del genoma del HDV,
- f) el fragmento R'2 que se extiende desde las posiciones 984 a 331 del genoma del HDV,
- g) el ADNc que codifica la proteína sHD, de secuencia SEC ID N°: 4,
- 20 h) el ADNc que codifica la proteína LHD de secuencia SEC ID N°: 2, y
- i) los cebadores de la secuencia SEC ID N°: 38 y SEC ID N°: 47

3. Método de detección de un HDV variante tal como se define en la reivindicación 1 o la reivindicación 2, por hibridación y/o amplificación, realizado a partir de una muestra biológica, cuyo

25 método se **caracteriza por que** comprende:

- (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 usando un par de cebadores seleccionado en el grupo constituido por los cebadores adecuados para amplificar una de las siguientes regiones del
- 30 ARN genómico del HDV: R0, R1, R2, R3, R'1 y R'2 y
- (3) el análisis del producto amplificado por comparación con la molécula de secuencia SEC ID N°: 1.

4. Método de detección de acuerdo con la reivindicación 3, **caracterizado por que** la etapa (3) de análisis puede realizar por restricción, secuenciación o hibridación

35 5. Método de acuerdo con la reivindicación 3 o la reivindicación 4, **caracterizado por que** los cebadores específicos para la amplificación de las regiones R0, R1, R2, R3, R'1 y R'2, usados en la etapa (2) se seleccionan del grupo constituido por:

- los cebadores 900S (SEC ID N°: 33) y 1280AS (SEC ID N°: 34), para la amplificación de R0 (aproximadamente 400 pb),
- los cebadores 320S (SEC ID N°: 39) y 1280AS (SEC ID N°: 34), para la amplificación del fragmento R1 (aproximadamente 960 pb),
- 5 - los cebadores 900S (SEC ID N°: 33) y 320AS (SEC ID N°: 45), para la amplificación de R2 (aproximadamente 1100 pb), que contiene el gen sHD correspondiente a las posiciones 1598-950,
- los cebadores 1480S (SEC ID N°: 46) y 440AS (SEC ID N°: 47), para la amplificación de R3 (aproximadamente 650 pb),
- los cebadores 318S (SEC ID N°: 35) y 1150AS (SEC ID N°: 36), para la amplificación de R'1
- 10 (aproximadamente 850 pb),
- los cebadores 960S (SEC ID N°: 37) y 345AS (SEC ID N°: 38), para la amplificación de R'2 (aproximadamente 1050 pb),

6. Método de de detección y de genotipado del HDV a partir de una muestra biológica, cuyo
15 método se **caracteriza por que 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 del HDV,
- (c) un primer tratamiento de moléculas de ácidos nucleicos amplificadas por las enzimas de restricción
- 20 *SmaI* y *XhoI*, para producir un primer conjunto de fragmentos de restricción,
- (d) un segundo tratamiento de moléculas de ácidos nucleicos por la enzima de restricción *SacII*, para producir un segundo conjunto de fragmentos de restricción
- (e) el análisis combinado, por RFLP (*Restriction Fragment Length Polymorphism (Polimorfismo de Longitud de Fragmentos de Restricción)*), de los dos conjuntos de fragmentos restricción producidos,
- 25 para detectar la presencia y/o determinar el tipo de HDV presente en dicha muestra biológica.

7. Método de acuerdo con la reivindicación 6, **caracterizado por que** la etapa (b) de amplificación, de las moléculas de ácido nucleico de dicha muestra, por RT-PCR se realiza ventajosamente con los cebadores 900S (SEC ID N°: 33) y 1280AS (SEC ID N°: 34).

30 8. Método de acuerdo con la reivindicación 6 o la reivindicación 7, **caracterizado por que** también comprende:

- (a) la amplificación de las moléculas de ácido nucleico de dicha muestra por RT-PCR con los cebadores 900S (SEC ID N°: 33) y 320AS (SEC ID N°: 45) para amplificar la región R2, y
- 35 (b) la secuenciación directa de la región R2 amplificada y la comparación con la molécula de ARN de la secuencia SEC ID N°: 1

9. Vector recombinante, particularmente un plásmido que comprende un inserto constituido por una molécula de ácido nucleico de acuerdo con la reivindicación 1 o la reivindicación 2.
10. Célula transformada por una molécula de ácido nucleico de acuerdo con la reivindicación 1 o la reivindicación 2.
11. Productos de traducción codificados por la molécula de ARN de secuencia SEC ID N°: 1 o por secuencias complementarias sentido o antisentido.
12. Proteínas, **caracterizadas por que** una de las moléculas de ácido de nucleico de acuerdo con la reivindicación 1 o la reivindicación 2, las codifican.
13. Proteína de acuerdo con la reivindicación 12, **caracterizada por que** se selecciona del grupo constituido por:
- la proteína LHD de dFr45, que presenta la secuencia SEC ID N°: 3, y
 - la proteína sHD de dFr45, que presenta la secuencia SEC ID N°: 5
14. Péptido, **caracterizado por que** está constituido por un fragmento de una proteína de acuerdo con la reivindicación 12 o la reivindicación 13, seleccionado del grupo constituido por:
- el péptido A constituido por los 19 aminoácidos del extremo carboxi-terminal de la secuencia SEC ID N°: 3,
 - el péptido B de secuencia (código de una letra) RLPLLECTPQ (SEC ID N°: 59) constituido por los 10 aminoácidos del extremo carboxi-terminal de la secuencia SEC ID N°: 3, y
 - el péptido C constituido por los 9 aminoácidos anteriores a la secuencia SEC ID N°: 59 (SEC ID N°: 60)
15. El uso de una molécula de ácido nucleico de acuerdo con la reivindicación 1 o la reivindicación 2 o de una proteína de acuerdo con una cualquiera de las reivindicaciones 11 a 14 para la preparación de un kit de detección y genotipado de un HDV.

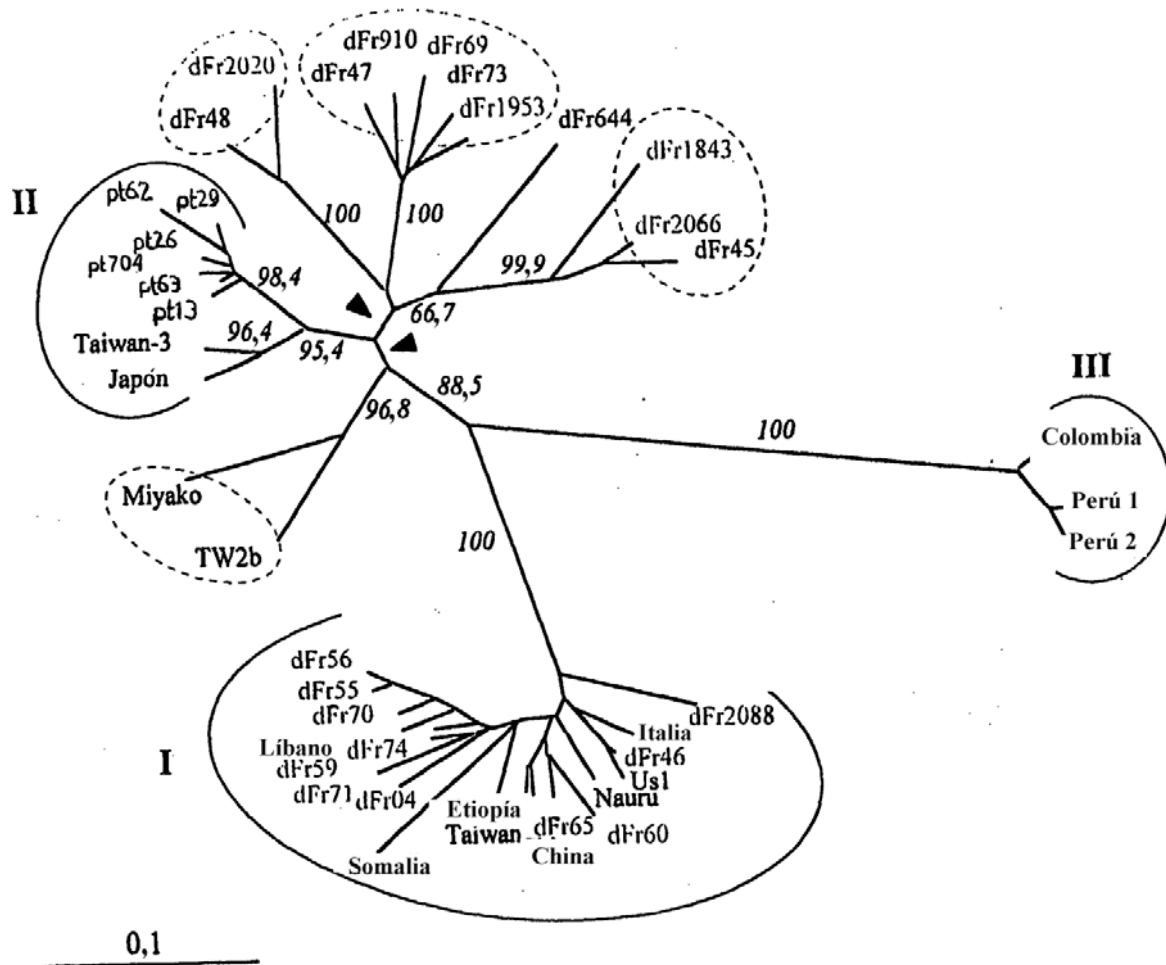


FIGURA 1

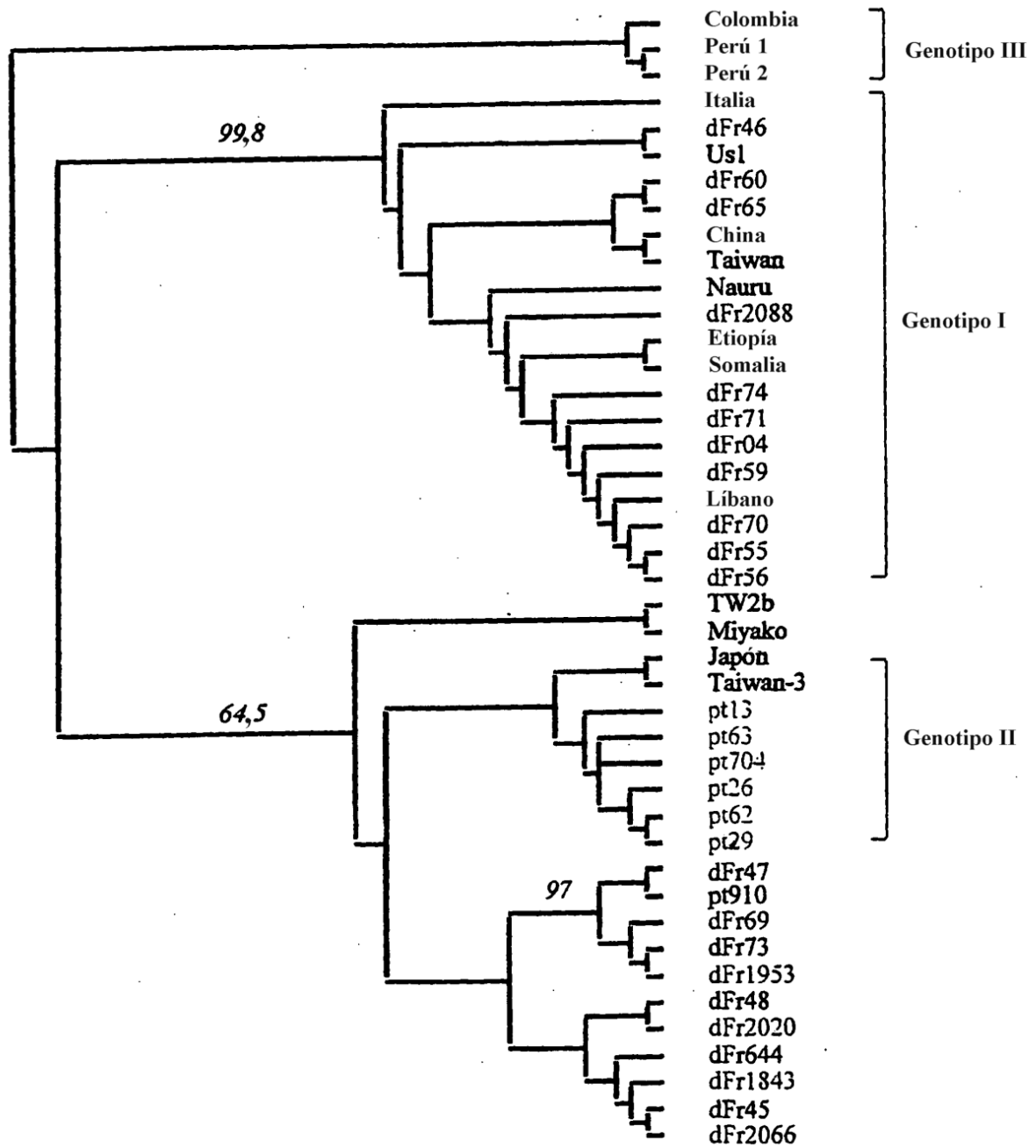


FIGURA 2

Pacientes		Marcadores HDV				Hígado	Tratamiento
		Ag-HD	IgM	IgG	ARN		
Nº de paciente/ anónimo	Sexo	Edad	Origen				
dFr45	F	43	Camerún	-	+	+	Anatomo- patología HCA IFN x 3 9M 12 meses
dFr47	F	38	Guinea	-	+	+	NT
dFr48	F	40	Pologne (estancias en Camerún)	-	+	+	IFN x 3 5M 6 meses
dFr73	M	32	Costa de Marfil	-	+	+	IFN x 3 10M 12 meses
dFr644	F	30	Congo	-	+	+	Transplante 09/00 cirrosis
dFr910	F	32	Mali	-	+	+	IFN x 3 3M 3 meses

FIGURA 3

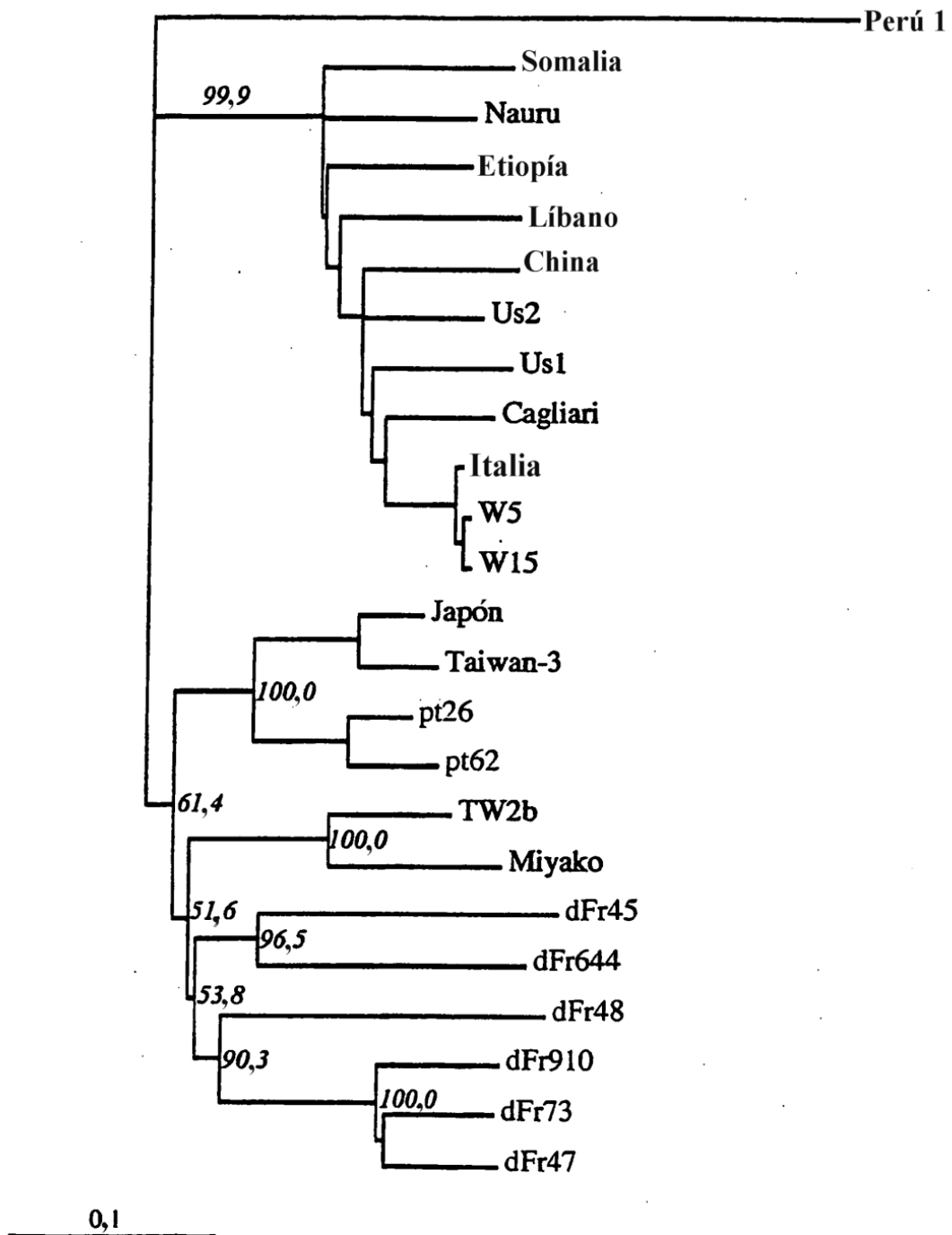


FIGURA 4

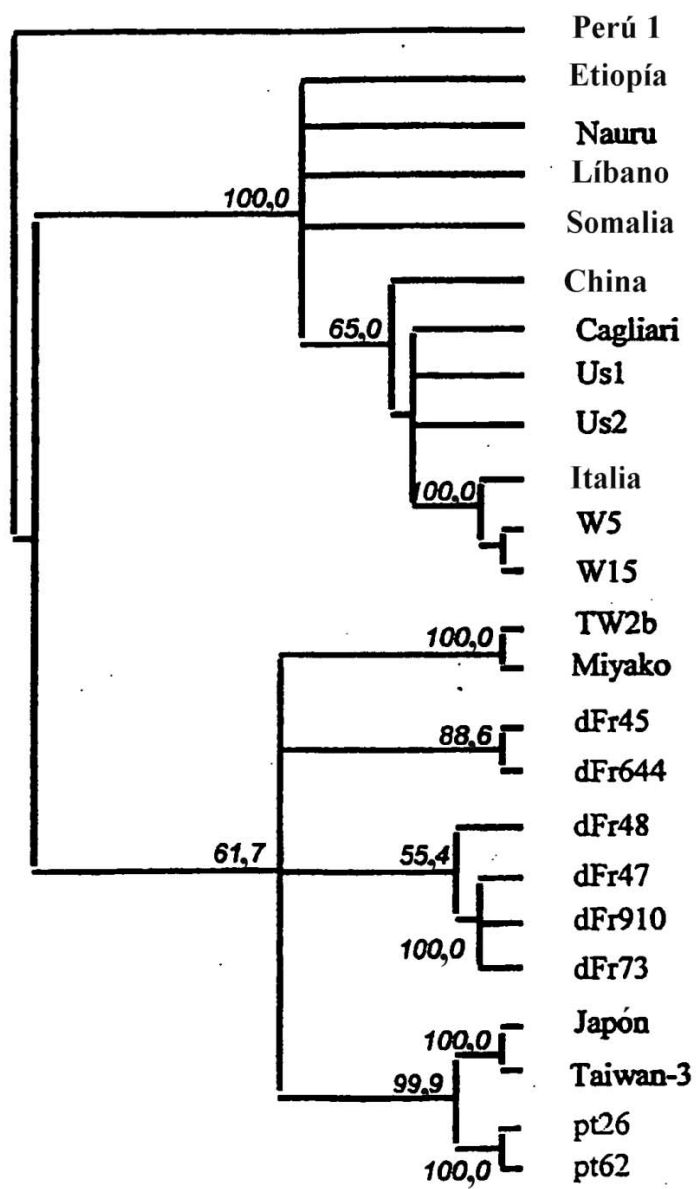


FIGURA 5

```

1 ERRVAGPRVGDVNPLDGGPRGAPGGGFVPSMHDIPESFFTRRGDGLDVRGAQEFZVSPQPPFRLPLLE-CTPQ
2 .....E.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.....-....
6 .....PG.S.....Q.TGV.....T.E...I..T.G..WVN.V..GQ.....-....
7 .....PE.....Q.L.V.....S.T.S.....N.Q..WVD.G.R.....-....
8 .....PGTS.....Q.LGV.....S.T.E.....N.QY.WVN.G.R.....-....
9 .....PQ.....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..WGSPST.....-....
12 K..T.....PG.D.....T.LHV.....S.T.E.....T.Q..WGNT.P.R.....-....
13 ...A..P.G...E..Q.....QGV.....H.E.....TGG..ZDILE.SD.PFSPQS-R..
14 .....P.G.I...E.....LQGV.....S.T.E...I..NRG..WDILE.AD.PFSPQS-R..
15 S..I..SA.G..E..S.....QG.....H..E.....T.TGG..ZDILE.SD.PFSPQS-R..
16 ...T..P.G...M..P.....LQGV.....S.T.E.I.I..T.Q..WYGFT...GYWVPG.TQ.
  ** * * * * * * * * : * * * * * * * * : * * * * *

```

FIGURA 6