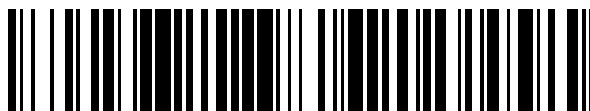


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 386 742**

51 Int. Cl.:
C07H 19/00 (2006.01)
C12Q 1/00 (2006.01)
C12Q 1/66 (2006.01)
C12Q 1/68 (2006.01)
C12Q 1/70 (2006.01)
G01N 33/53 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 96 Número de solicitud europea: **08000552 .3**
96 Fecha de presentación: **11.03.2003**
97 Número de publicación de la solicitud: **1947104**
97 Fecha de publicación de la solicitud: **23.07.2008**

54 Título: **Método y composiciones para identificar y caracterizar la Hepatitis C**

30 Prioridad:
11.03.2002 US 363603 P

45 Fecha de publicación de la mención BOPI:
28.08.2012

45 Fecha de la publicación del folleto de la patente:
28.08.2012

73 Titular/es:
**LAB 21 LIMITED
184 CAMBRIDGE SCIENCE PARK
CAMBRIDGE CB4 0GA, GB**

72 Inventor/es:
Holland-Staley, Carol

74 Agente/Representante:
de Elzaburu Márquez, Alberto

ES 2 386 742 T3

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

DESCRIPCIÓN

Método y composiciones para identificar y caracterizar la hepatitis C

Solicitudes relacionadas

5 La presente solicitud reivindica prioridad de la solicitud de patente provisional de EE.UU. nº de serie 60/363.603 presentada el 11 de marzo de 2002.

Fundamento de la invención

El virus de la hepatitis C (en lo sucesivo abreviadamente HCV, por sus iniciales en inglés) es el principal agente etiológico para la hepatitis no-A, no-B. Se estima que están infectadas alrededor de 400 millones de personas o más del 2% de la población mundial (Di Bisceglie, A. M. (1998), *Hepatitis C*. Lancet 351: 351-5; Houghton M. (1996), pp. 1035-1058. En B. N. Fields y D. M. Knipe y H. P. M. (ed.), *Fields' Virology*, 3ª ed. Lippincott-Raven, Philadelphia/New York). La enfermedad por HCV habitualmente es el resultado de una infección crónica en el 60 a 80% de los individuos infectados, de los cuales el 20% progresa a cirrosis, carcinoma hepatocelular o insuficiencia hepática crónica. En todo el mundo se han identificado al menos 6 genotipos virales principales y se han propuesto más de 50 subtipos de HCV (Simmonds, P. et al., (1994) Identification of genotypes of hepatitis C virus by sequence comparisons in the core, E1 and NS-5 regions. *J. Gen. Virol.* 75: 1053-1061). Estos genotipos están basados en una diversidad de secuencias de nucleótidos y aminoácidos, difiriendo los aislados más divergentes en más del 30% (Bréchet, C. et al., (1996) *Hepatitis C virus: Molecular biology and genetic variability*. *Digestive Diseases and Sciences* 41: 6S-21S; Bukh, J. et al., (1995) Genetic heterogeneity of hepatitis C virus: quasispecies and genotypes. *Semin. Liver Dis.* 15: 41-63). Entre los diferentes tipos, hay regiones que están altamente conservadas y tienen una homología de secuencias próxima al 100%, sin embargo, en las regiones altamente variables, tales como las proteínas de la envoltura es <70% (Booth, J. C. et al., (1998) Comparison of the rate of sequence variation in the hypervariable region of E2/NS1 region of hepatitis C virus in normal and hypogammaglobulinemics patients. *Hepatology* 27: 223-27; Hayashi, N. et al., (1993) Molecular cloning and heterogeneity of the human hepatitis C virus (HCV) genome. *J. Hepatol.* 17: S94-107; Hijikata, M. et al., (1991) Hypervariable regions in the putative glycoprotein of hepatitis C virus. *Biochem. Biophys. Res. Commun.* 175: 220-228; Kato, N. (1992) Marked sequence diversity in the putative envelope protein of hepatitis C viruses. *Virus Res.* 22: 107-123; Lesniewski, R. R. et al., (1993) Hypervariable 5'-terminus of hepatitis C virus E2/NS1 encodes antigenically distinct variants. *J. Med. Virol.* 40: 150-156; Pozzetto, B. et al., (1996) Structure, genomic organization, replication and variability of hepatitis C virus. *Nephrol. Dial. Transplant.* 11 [Suppl 4]: 2-5; Vizmanos, J. L. et al., (1998), Degree and distribution of variability in the 5'-untranslated, E1, E2/NS1 and NS5 regions of the hepatitis C virus (HCV). *J. Viral Hepat.* 5: 227-240).

El HCV es un miembro de la familia Flaviviridae cuyos otros miembros incluyen Flavivirus y Pestivirus (Kato, N. et al., (1991) Molecular structure of the Japanese hepatitis C viral genome. *FEBS Lett.* 280: 325-328). El genoma del HCV es RNA de cadena positiva de de aproximadamente 9,5 kb que contiene un único marco de lectura abierto que codifica una poliproteína de 3010 a 3033 aminoácidos (Takamizawa, A. et al., (1991) Structure and organization of the hepatitis C virus genome isolated from human carriers. *J. Virol.* 65: 1105-1113). El procesamiento proteolítico de la poliproteína se realiza por proteasas del hospedante y virales. Las peptidasas del hospedante escinden las proteínas estructurales que están localizadas en el extremo 5', mientras que las proteasas virales escinden las proteínas no estructurales. Estas escisiones dan como resultado al menos 10 proteínas virales en el orden NH2-C-E1-E2-p7-NS2-NS3-NS4A-NS4B-NS5A-NS5B-COOH. También existen regiones no codificadoras localizadas en los extremos 5' y 3' que están implicadas en la unión al ribosoma y en la iniciación de la replicación, respectivamente.

Una de las proteasas virales, la proteasa NS3, está codificada por la región N-terminal del gen NS3 del HCV. Consiste en 181 aminoácidos y es una serina-proteasa similar a quimotripsina responsable de la escisión de las proteínas no estructurales del HCV (Bartenschlager, R. et al., (1993) Nonstructural protein 3 of the hepatitis C virus encodes a serine-type protease required for cleavage at the NS3/4 and NS4/5 junctions. *J. Virol.* 67: 3835-3844; Eckart, M. R. et al., (1993). The hepatitis C virus encodes a serine protease involved in processing of the putative nonstructural proteins from the viral polyprotein precursor. *Biochem. Biophys. Res. Commun.* 192: 399-406; Gallinari, P. et al., (1998) Multiple enzymatic activities associated with recombinant NS3 protein of hepatitis C virus. *J. Virol.* 72: 6758-6769; Hahm, B. et al., (1995) NS3-4A of hepatitis C virus is a chymotrypsin-like protease. *J. Virol.* 69: 2534-2539; Hijikata, M. et al., (1993) Two distinct proteinase activities required for the processing of a putative nonstructural precursor protein of hepatitis C virus. *J. Virol.* 67: 4665-4675; Hijikata, M. et al., (1993) Proteolytic processing and membrane association of putative nonstructural proteins of hepatitis C virus. *Proc. Natl. Acad. Sci. USA.* 90: 10773-10777; Manabe, S. et al., (1994) Production of nonstructural proteins of hepatitis C virus requires a putative viral protease encoded by NS3. *Virology* 198: 636-644; Tomei, L. et al., (1993) NS3 is a serine protease required for processing of hepatitis C virus polyprotein. *J. Virol.* 67: 4017-4026). El procesamiento de las proteínas estructurales por NS3 ocurre en una cascada bien ordenada ocurriendo la primera escisión entre NS3 y NS4A seguida por NS5A-NS5B, NS4A-NS4B y NS4B-NS5A (Bartenschlager, R. et al., (1994) Kinetic and structural analysis of hepatitis C virus polyprotein processing. *J. Virol.* 68: 5045-5055; D'Souza, E. D. et al., (1994) Analysis of NS3-mediated processing of the hepatitis C virus non-structural region in vitro. *J. Gen. Virol.* 75: 3469-3476; Eckart M. R. et al., 1993, supra; Failla, C. et al., (1994) Both NS3 and NS4A are required for proteolytic processing of

hepatitis C virus nonstructural proteins. *J. Virol.* 68: 3753-3760; Kolykhalov, A. A. et al., (1994) Specificity of the hepatitis C virus NS3 serine protease: effects of substitutions at the 3/4A, 4A/4B, 4B/5A, and 5A/5B cleavage sites on polyprotein processing. *J. Virol.* 68: 7525-7533; Shimotohno, K. et al., (1995) Processing of the hepatitis C virus precursor protein. *J. Hepatol.* 22: 87-92; Shoji, I. et al., 1999 Internal processing of hepatitis C virus NS3 protein. *Virology* 254: 315-323; Tomei, L. et al., 1993, supra). La escisión entre NS3 y NS4A ocurre en cis, mientras que las otras escisiones son en trans. El co-factor codificado por el virus, NS4A, es necesario para la función eficaz de NS3. La eficacia del procesamiento proteolítico se ha demostrado que aumenta espectacularmente en presencia de la proteína NS4A (Failla, C. et al., (1995) An amino-terminal domain of the hepatitis C virus NS3 proteinase is essential for the interaction with NS4A. *J. Virol.* 69: 1769-1777; Failla, 1994, supra; Gallinari, P. et al., 1999 Modulation of hepatitis C virus NS3 protease and helicase activities through the interaction with NS4A. *Biochemistry* 38: 5620-5632; Lin, C. et al., (1995) A central region in the hepatitis C virus NS4A protein allows formation of an active NS3-NS4A serine protease complex in vivo and in vitro. *J. Virol.* 69: 4373-4380; Satoh, S. et al., (1995) The N-terminal region of hepatitis C virus nonstructural protein 3 (NS3) is essential for stable complex formation with NS4A. *J. Virol.* 69: 4255-4260; Tanji, Y. et al., (1995) Hepatitis C virus-encoded nonstructural protein NS4A has versatile functions in viral protein processing. *J. Virol.* 69: 1575-1581). Además, NS3 contiene un átomo de zinc unido tetraédricamente, que parece desempeñar una función estructural (De Francesco, R. et al., (1996) A zinc binding site in viral serine proteinases. *Biochemistry* 35: 13282-13287; Stempniak, M. et al., (1997) The NS3 proteinase domain of hepatitis C virus is a zinc-containing enzyme. *J. Virol.* 71: 2881-2886).

Recientemente se ha logrado un progreso considerable en determinar cómo la proteasa NS3 procesa el polipéptido de HCV (Kolykhalov, A. A. et al., 1994 supra; Steinkühler, C. et al., (1996) Activity of purified hepatitis C virus protease NS3 on peptide substrates. *J. Biol. Chem.* 271: 6694-6700; Urbani, A. et al., (1997). Substrate specificity of the hepatitis C virus serine protease NS3. *J. Biol. Chem.* 272: 9204-9209). Los modelos de cómo la proteasa interactúa con los co-factores y el sustrato han identificado cuatro dominios, que están implicados en la función enzimática. Barbato, G. et al., 1999 The solution structure of the N-terminal proteinase domain of the hepatitis C virus (HCV) NS3 protein provides new insights into its activation and catalytic mechanism. *J. Mol. Biol.* 289: 371-384. Estas son la triada catalítica, el co-factor y los sitios de unión al metal y la cavidad de unión al sustrato. Estos dominios están bien definidos y contienen residuos de aminoácidos que está altamente conservados en los genes de las proteasas de HCV secuenciados hasta la fecha (Véase la Figura 1 en Holland-Staley, C. A., et al., (2002) Genetic diversity and response to IFN of the NS3 protease gene from clinical strains of the hepatitis C virus. *Arch. Virol.* 147:1385-1406, que se incorpora en la presente memoria como referencia). A pesar de este reciente estallido de información estructural y estudios que muestran la implicación directa y la conservación de residuos de aminoácidos el impacto de la variabilidad de la secuencia natural sobre la función enzimática no está bien entendido. Además se desconocen los efectos de la terapia anti-viral sobre la secuencia de la proteasa de NS3 del HCV. La elucidación de la diversidad genética natural de la proteasa NS3 del HCV en muestras de pacientes tiene tanto significado médico como interés teórico. Aunque todas las proteasas NS3 del HCV secuenciadas hasta la fecha contienen aminoácidos del sitio activo conservados, la variación de secuencia a través de todo el gen NS3 es significativa (Martell, M. et al., (1992) Hepatitis C virus (HCV) circulates as a population of different but closely related genomes: quasispecies nature of HCV distribution. *J. Virol.* 66: 3225-32; Okamoto, H. et al., (1992) Genetic drift of hepatitis C virus during an 8.2-year infection in a chimpanzee: variability and stability. *Virology* 190: 894-899. También la presencia de múltiples especie dentro del mismo paciente, conocidas como cuasi-especies, crea un problema potencial para el desarrollo y resistencia de fármacos. Entender la magnitud de la variación de la secuencia de NS3 en cepas clínicas permitirá un desarrollo más eficaz de fármacos que tienen como diana la proteasa del HCV. Para esto deben estar disponibles datos que describen la variabilidad de la secuencia como ocurre en las personas infectadas con HCV.

Las helicasas son enzimas que son responsables del desenrollado de los dúplex DNA/DNA, RNA/DNA y RNA/RNA en una dirección 3' a 5'. (Bartenschlager, R. et al., 1993, supra; Hahm B. et al., 1995, supra; Tomei et al., 1993,). Además, se ha propuesto que las enzimas helicasas desempeñan funciones en la replicación viral y la recombinación, el control viral de las funciones celulares del hospedante, la estabilidad del mRNA incluyendo el empalme o procesamiento, la transcripción, el transporte, y la iniciación de la traducción del RNA (Luking, et al., (1998). The protein family of RNA helicases *Crit. Rev. Biochem. Mol. Biol.* 33: 259-296). La helicasa/NTPasa de NS3 de HCV contiene 450 aminoácidos. El componente NTPasa hidroliza los 5'-trifosfatos de nucleósidos, que proporcionan los requisitos energéticos. La helicasa/NTPasa, junto con la proteasa NS3 y la polimerasa NS5b dependiente de RNA, se cree que forman un complejo grande, que es responsable de la replicación del RNA viral. Recientemente, se ha conseguido un progreso considerable en determinar la estructura de la helicasa NS3 del HCV y su mecanismo de desenrollado de los dúplex. Al menos se han publicado tres estructuras cristalinas diferentes (Paolini, Cho et al., 1998, supra; Kim 1998 et al., supra; y Yao et al., 1997, supra). En todas las tres, la enzima se demostró que contenían 3 dominios, siendo los dominios 1 y 2 estructuralmente similares. La enzima contiene siete restos, que incluyen dos restos que están implicados en la unión y la hidrólisis de NTP (el resto I [GxGKS] y el resto II [DExH]) y uno que está implicado en la hidrólisis de ATP y el desenrollado del RNA (resto VI). La mayoría de los restos están localizados entre los dominios estructurales 1 y 2, con el dominio 3 separado de los otros dos por la unión del nucleótido. Los investigadores han identificado residuos altamente conservados dentro de cada resto. Estos restos, junto con los análisis de comparación con las otras helicasas, revelan que son un miembro de la familia DEAD-box de helicasas de RNA, específicamente la subfamilia DExH. Se requiere magnesio para ambas actividades de helicasa y NTPasa.

La única terapia aprobada por la FDA para la infección de la hepatitis C es interferón (IFN) o interferón pegilado con o sin ribavirina. La producción de interferón ha demostrado ser inducida por bacterias, parásitos y virus, así como en respuesta a los tumores. Los interferones son proteínas secretadas de la familia de las citoquinas, que inhiben indirectamente el ciclo de la vida viral uniéndose a los receptores celulares, induciendo de este modo la síntesis de proteínas (Hijikata, M. et al., (1993) Proteolytic processing and membrane association of putative nonstructural proteins of hepatitis C virus. Proc. Natl. Acad. Sci. USA. 90: 10773-10777). Los genes inducibles por el interferón contienen una región promotora, denominada elemento de respuesta estimulado por IFN (en lo sucesivo abreviadamente ISRE, por la expresión inglesa IFN-Stimulated Response Element). Más de 30 genes son inducidos por interferón, sin embargo, se desconoce la función de la mayoría de estos genes (Holmes, E. C. et al., (2000) The causes and consequences of genetic variation in dengue virus. Trends Microbiol. 8: 74-77).

Recientemente se ha sugerido que un corto tramo de 40 aminoácidos en el gen NS5A de HCV desempeña una función en la resistencia al IFN, sin embargo, aunque esto parece ser cierto en algunos aislados japoneses otros investigadores tienen datos en conflicto. Los efectos de la terapia del interferón sobre los otros genes estructurales son desconocidos.

15 Sumario de la invención

La presente invención se basa al menos parte, en el descubrimiento de nuevas secuencias de ácidos nucleicos aisladas para amplificar y secuenciar porciones del genoma del HCV. La presente invención proporciona nuevas secuencias y métodos de usarlas, que amplifican consistentemente el genoma del HCV. Las secuencias de ácidos nucleicos de la presente invención son útiles como reactivos para detectar e identificar secuencias virales en muestras biológicas y facilitar más caracterización del genoma de HCV. Los métodos descritos en la presente invención, las secuencias de ácidos nucleicos descritos como la SEC ID NO: 3; y sus equivalentes funcionales, así como kits que las contienen, son útiles para determinar con precisión la presencia del HCV en muestras biológicas, así como la secuencia específica de una porción del genoma del HCV. Las secuencias de ácidos nucleicos y los métodos de la presente invención pueden ser útiles para: (i) detectar la infección por HCV, incluyendo la detección en etapas tempranas; (ii) la identificación del tipo de infección por HCV; (iii) la detección de las cepas variantes del HCV incluyendo la heterogeneidad en un paciente; (iv) la detección de una mutación en el ácido nucleico del HCV que es responsable de la resistencia o sensibilidad a una terapia; (v) la detección de nuevas mutaciones en el genoma del HCV que están correlacionadas con la resistencia o sensibilidad a una terapia con fármacos; (vi) determinar si será eficaz el tratamiento con un agente, por ejemplo, un agente anti-viral; (vii) determinar si el tratamiento con un agente, por ejemplo, un agente anti-viral, debe ser o no continuado; (viii) la identificación de la interacción entre el HCV y otros virus y/o enfermedades; (ix) generar un ácido nucleico, por ejemplo, una vacuna de DNA, que puede ser específica del paciente; y (x) desarrollo de nuevos fármacos, por ejemplo, basado en cristalografía de rayos X.

La presente invención se refiere a nuevas secuencias de cebadores aisladas del genoma del HCV descritas en la presente memoria como la SEC ID NO: 3 (Tablas 1A-1G). Las secuencias que corresponden a los subtipos 1a y 1b del HCV son las descritas como las SEC ID NO: 1-23 y la SEC ID NO: 64 (Tablas 1A y 1G); las secuencias que corresponden al subtipo 2a del HCV son las descritas como las SEC ID NO: 24-27 (Tabla 1B); las secuencias que corresponden al subtipo 2b del HCV son las descritas como las SEC ID NO: 28-42 (Tabla 1C); las secuencias que corresponden al subtipo 3a del HCV son las descritas como las SEC ID NO: 43-46 (Tabla 1D); secuencias que corresponden al subtipo 3b del HCV son las descritas como las SEC ID NO: 47-50 (Tabla 1E); y las secuencias que corresponden al subtipo 4a del HCV son las descritas como las SEC ID NO: 51-63 (Tabla 1F).

En una realización, la presente invención se refiere a un oligonucleótido descrito como SEC ID NO: 3. En otra realización, el oligonucleótido de la invención es al menos 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o más idénticos a la secuencia de nucleótido descrita como SEC ID NO: 3. En otra realización, el oligonucleótido de la SEC ID NO: 3 tiene una longitud de al menos 4, 8, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 o más nucleótidos. Se apreciará que el extremo 5' puede contener la mayor variabilidad (por ejemplo, sustituciones de ácidos nucleicos), permaneciendo todavía funcional (por ejemplo, ser capaz de hibridarse a la porción de los genes NS3 y NS4 del genoma del HCV). En otra realización, la invención proporciona una combinación de uno o más oligonucleótidos de la presente invención. En todavía otra realización, la invención proporciona un conjunto de oligonucleótidos, también denominados en la presente memoria "cebadores" y "secuencias de cebadores de ácidos nucleicos" seleccionadas del grupo que consiste en dos o más de los oligonucleótidos de la presente invención. En incluso otra realización, la invención proporciona oligonucleótidos que son capaces de amplificar una muestra de un paciente que tiene una secuencia de nucleótidos descrita como la SEC ID NO: 3, o uno de sus complementos. En un aspecto de la invención, los oligonucleótidos de la presente invención comprenden un marcador para detección. Dichos marcadores pueden ser por ejemplo, marcadores radiactivos que pueden ser incorporados por métodos conocidos (por ejemplo, traslado de muescas o tratamiento con quinasas), isótopos radioactivos, biotina, grupos fluorescentes, grupos quimioluminiscentes (por ejemplo, dioxetanos, particularmente dioxetanos activados), digoxigenina, enzimas, anticuerpos, agentes luminiscentes, agentes precipitantes, colorantes, sus combinaciones y similares. Otro aspecto de la invención comprende un vector que contiene un oligonucleótido de la presente invención.

En otra realización, la invención proporciona un oligonucleótido para amplificar una secuencia de ácido nucleico contenida en una muestra, en donde la muestra contiene una secuencia de nucleótidos descrita como la SEC ID NO: 3, o uno de sus complementos. En un aspecto de la invención, la muestra es de un paciente infectado con HCV.

5 En otra realización, la invención proporciona una secuencia de ácido nucleico, en donde la secuencia de ácido nucleico es un promotor-cebador que comprende un oligonucleótido of la presente invención, en donde una porción 5' de la secuencia incluye una secuencia de promotor.

10 Las secuencias de ácido nucleico amplificado del HCV generadas por los métodos descritos en la presente memoria pueden ser clonadas en un vector hospedante con un promotor de expresión y ser usadas como una vacuna de DNA para amplificar una respuesta del paciente al virus HCV, proporcionando de este modo una vacuna de DNA para el HCV específica del paciente. En otro aspecto de la invención, el clon puede usarse para generar RNA para uso como una vacuna antisentido.

15 Se apreciará que una muestra usada en los métodos de la presente invención puede ser una muestra biológica, por ejemplo, de un paciente infectado con HCV. Dichas muestras pueden incluir, sin limitación, sangre, plasma, suero, fluido de la médula espinal, fluido linfático, secciones externas de la piel, los tractos respiratorio, intestinal, y genitourinario, lágrimas, saliva, leche, orina, células de la sangre, tumores, órganos, DNA genómico, RNA, o cDNA en solución o unido a un sustrato, y también muestras de constituyentes de cultivos celulares in vitro incluyendo, pero sin limitación, medio acondicionado resultante del crecimiento de células en un medio de cultivo, células infectadas supuestamente por virus, células recombinantes, y componentes celulares, por ejemplo, cromosoma(s), orgánulos, tejido incrustado en parafina o membranas aisladas de una célula.

20 En una realización de la invención, se analiza la porción de serina-proteasa del gen NS3 del HCV. Esto puede conseguirse amplificando y secuenciando la porción de serina-proteasa del gen con las composiciones y métodos de la presente invención. Cuando se emplea un par de cebadores como cebadores de secuenciación, uno o ambos de los miembros del par puede(n) estar adecuadamente marcado(s), por ejemplo, con marcadores radiactivos que pueden ser incorporados por métodos conocidos (por ejemplo, traslado de muesca o tratamiento con quinasa),
25 isótopos radiactivos, biotina, grupos fluorescentes, grupos quimioluminiscentes (por ejemplo, dioxetanos, particularmente dioxetanos activados), digoxigenina, enzimas, anticuerpos, agentes luminiscentes, agentes precipitantes, colorantes, sus combinaciones y similares.

30 En otra realización de la invención, se analiza la porción de helicasa del gen NS3 del HCV gene. Esto puede conseguirse amplificando y secuenciando la porción de helicasa del gen con las composiciones y métodos de la presente invención. Cuando se emplea un par de cebadores como cebadores de secuenciación, uno o ambos de los miembros del par puede(n) estar adecuadamente marcado(s), por ejemplo, con marcadores radiactivos que pueden ser incorporados por métodos conocidos (por ejemplo, traslado de muesca o tratamiento con quinasa), isótopos radiactivos, biotina, grupos fluorescentes, grupos quimioluminiscentes (por ejemplo, dioxetanos, particularmente dioxetanos activados), digoxigenina, enzimas, anticuerpos, agentes luminiscentes, agentes precipitantes, colorantes,
35 sus combinaciones y similares.

40 Las combinaciones de cebadores, por ejemplo, que incluyen al menos un cebador directo y un cebador inverso, que juntos pueden usarse para la amplificación y/o secuenciación de una porción relevante del genoma del HCV, por ejemplo, el gen NS3 o NS4, pueden estar adecuadamente empaquetadas en un kit. Los pares anidados de cebadores de amplificación y secuenciación son los preferidos. En una realización, la invención proporciona un kit para la detección del HCV, que comprende uno o más oligonucleótidos de la presente invención. En otra realización, el kit comprende un conjunto de cebadores seleccionados del grupo que consiste en los oligonucleótidos de la presente invención. Los cebadores de dichos kits pueden estar marcados o no marcados. El kit puede incluir también reactivos adicionales, tales como reactivos para realizar una reacción en cadena de la polimerasa (PCR), una transcriptasa inversa para la conversión del RNA del HCV en cDNA para amplificación, DNA-polimerasas,
45 materiales de alimentación de dNTP y ddNTP. El kit también puede incluir instrucciones para su uso.

50 La presente invención también se refiere a nuevos métodos para amplificar un ácido nucleico diana, por ejemplo, regiones específicas del genoma del HCV de un paciente, tal como los genes NS3 y NS4 del genoma del HCV, usando los oligonucleótidos de la presente invención. Los nuevos métodos de la presente invención pueden ser usados en, por ejemplo: (i) detectar una infección por HCV, incluyendo la detección en etapas tempranas; (ii) la identificación del tipo de infección por HCV; (iii) la detección de cepas variantes del HCV incluyendo heterogeneidad en un solo paciente; (iv) la detección de una mutación en el ácido nucleico del HCV que es responsable de la resistencia o sensibilidad a una terapia; (v) la detección de nuevas mutaciones en el genoma del HCV que están correlacionadas con la resistencia o sensibilidad a la terapia; (vi) determinar si será eficaz el tratamiento con un agente, por ejemplo, un agente antiviral; (vii) determinar si debe o no deber ser continuado el tratamiento con un agente, por ejemplo, un agente antiviral; (viii) identificación de la interacción entre el HCV y otros virus y/o enfermedades; (ix) generar un ácido nucleico, por ejemplo, DNA, vacuna, que puede ser específica para el paciente; y (x) desarrollo de nuevos fármacos, por ejemplo, basados en cristalografía de rayos X. Este análisis puede ser realizado por secuenciación directa, o usando otra técnicas para la caracterización de polimorfismos de secuencias, incluyendo pero sin limitación, hibridación con sondas de oligonucleótidos específicas de secuencias.

Por consiguiente, en una realización, la invención proporciona métodos para amplificar un ácido nucleico diana, por ejemplo, porciones del gen NS3 o NS4, combinando un ácido nucleico diana en condiciones que permitan que ocurra una reacción de amplificación con una o más secuencias de cebadores de ácidos nucleicos de la presente invención y otros agentes de amplificación necesarios, tales como una polimerasa de ácido nucleico y una pluralidad de nucleótidos. En un aspecto de la presente invención, la amplificación de una secuencia diana de ácido nucleico da como resultado una cantidad aumentada del ácido nucleico diana amplificado. En incluso otro aspecto de la invención, la secuencia de ácido nucleico diana es del genoma del HCV, por ejemplo, el gen NS3 o NS4A. En otros aspectos de la invención, el ácido nucleico diana es la porción de serina-proteasa del gen NS3. En todavía otros aspectos de la invención, el ácido nucleico diana es la porción de helicasa del gen NS3. En otro aspecto de la invención, se usa una polimerasa de ácido nucleico y puede seleccionarse del grupo que consiste en transcriptasa inversa y DNA-polimerasa termoestable. En otras realizaciones de la invención se secuencia el ácido nucleico amplificado de HCV. En todavía otro aspecto de la invención, la secuencia se evalúa respecto a mutaciones.

En otra realización de la presente invención se proporciona un método para detectar y caracterizar ácidos nucleicos de HCV en una muestra, que comprende poner en contacto la muestra con una o más secuencias de cebadores de ácidos nucleicos de la presente invención en condiciones tales que los ácidos nucleicos del HCV puedan hibridarse con los cebadores de la invención, someter a transcripción inversa y amplificar los ácidos nucleicos para obtener los ácidos nucleicos del HCV amplificados y detectar la presencia de los ácidos nucleicos del HCV amplificados. Los ácidos nucleicos del HCV amplificados pueden ser caracterizados además, por ejemplo, secuenciando los ácidos nucleicos amplificados. En un aspecto de la invención, la detección de la presencia de los ácidos nucleicos del HCV amplificados comprende poner en contacto los ácidos nucleicos del HCV amplificados con un oligonucleótido marcado para obtener ácidos nucleicos del HCV marcados e identificar los ácidos nucleicos marcados. En todavía otros aspectos de la invención, la amplificación de los ácidos nucleicos puede conseguirse por amplificación basada en la secuencia de ácidos nucleicos (NASBA), amplificación mediada por la transcripción (TMA), reacción en cadena de la polimerasa (PCR), otros métodos de amplificación de dianas, métodos de amplificación de señal o métodos de amplificación de sondas, tales como reacción en cadena de la ligasa. En todavía otro aspecto de la invención, los métodos comprenden además la etapa de secuenciar los ácidos nucleicos del HCV amplificados. En otra realización más, los métodos comprenden además evaluar la secuencia respecto a mutaciones.

Una vez secuenciados, las mutaciones en el genoma del HCV pueden ser identificadas comparando la secuencia con una secuencia conocida o de control, por ejemplo, una secuencia de consenso. Por tanto, la presente invención proporciona composiciones y métodos para identificar mutaciones particulares en el genoma del HCV. Con el conocimiento de una mutación particular en el genoma del HCV de una muestra de un paciente, puede ser identificada y administrada la terapia o terapia de combinación para el paciente. Por ejemplo, los métodos de la presente invención pueden ser usados para determinar un régimen de tratamiento particular dependiendo del genotipo del HCV que ha infectado al paciente. En otra realización, la invención proporciona composiciones y métodos para determinar si un paciente es sensible o resistente a un agente obteniendo una muestra de un paciente que comprende DNA, realizando la amplificación, por ejemplo, PCR, sobre el DNA de la muestra del paciente usando una secuencia de cebador de ácido nucleico de la presente invención, y analizando el producto amplificado, por ejemplo, por secuenciación, para identificar mutaciones particulares. En todavía otros aspectos de la invención, el agente puede ser un agente antiviral, tal como un inhibidor de proteasa, un inhibidor de la porción de serina-proteasa del gen NS3, un inhibidor de la porción de helicasa del gen NS3, alfa-interferón, ribavirina, interferón pegilado o una de sus combinaciones.

Debido a que la secuencia del cebador de la presente invención se derivó de regiones conservadas del genoma del HCV, las composiciones y métodos descritos en la presente memoria serán útiles en la detección y/o identificación de cepas variantes del HCV. Esto, a su vez, conducirá al desarrollo de agentes inmunológicos adicionales para la detección y diagnóstico de HCV, así como al desarrollo de agentes polinucleotídicos adicionales para la detección y/o el tratamiento del HCV.

En todavía otra realización, la invención proporciona un método para determinar si un agente puede ser usado o no pueden ser usado para tratar un paciente infectado con HCV obteniendo una muestra de un paciente infectado con el HCV, amplificando la muestra del paciente usando las secuencias de cebadores de ácidos nucleicos de la presente invención, secuenciando las secuencias de cebadores de ácidos nucleicos, amplificando e identificando las mutaciones en la secuencia del ácido nucleico del HCV amplificado que están correlacionadas con la resistencia o la sensibilidad a un agente, determinando de este modo si el agente puede ser usado o no puede ser usado para tratar un paciente infectado con HCV. En un aspecto de la invención, las secuencias de ácidos nucleicos son un conjunto de cebadores.

En otra realización, la invención proporciona un método para determinar si el tratamiento con un agente debe ser continuado en un paciente infectado con el HCV, comprendiendo el método obtener dos o más muestras que comprenden DNA de un paciente durante el curso del tratamiento, amplificar y secuenciar una secuencia de ácido nucleico diana de la muestra usando las secuencias de cebadores de ácidos nucleicos de la presente invención en los métodos descritos en la presente memoria, identificar las mutaciones presentes en la secuencia de ácido nucleico amplificado que están correlacionadas con la resistencia o la sensibilidad a un agente, y continuar el

tratamiento cuando las mutaciones identificadas en el producto amplificado no cambian durante el curso del tratamiento.

En todavía otra realización, la invención proporciona un método para determinar si el tratamiento con un agente no debe ser continuado en un paciente infectado con el HCV, comprendiendo el método obtener dos o más muestras que comprenden DNA de un paciente durante el curso del tratamiento, amplificar y secuenciar las secuencias de ácidos nucleicos dianas de las muestras usando las secuencias de cebadores de ácidos nucleicos de la presente invención en los métodos descritos en la presente memoria, identificar las mutaciones presentes en la secuencia de ácido nucleico amplificada que está correlacionadas con la resistencia o la sensibilidad a un agente, e interrumpir el tratamiento cuando cambian las mutaciones identificadas en el producto amplificado, por ejemplo, un aumento en el número de mutaciones durante el curso del tratamiento es un indicador de la resistencia a un agente.

Los Ejemplos específicos que se exponen más adelante también establecen una correlación entre el tipo de HCV, particularmente como se determina por análisis del gen NS3, y la probabilidad de que el virus sea sensible a la terapia antiviral, por ejemplo, la terapia con alfa-IFN. Por tanto, un aspecto adicional de la invención es un método para evaluar el estado de un paciente diagnosticado, o sospechoso de tener una infección por HCV, para determinar si la terapia, por ejemplo, la terapia con alfa-IFN, es probable que tenga éxito, que comprende las etapas de obtener una muestra del paciente que contenga DNA, comparar la porción de serina-proteasa del gen NS3 y/o la porción de helicasa del gen NS3, con secuencias del subtipo de consenso, en donde la presencia de una o más mutaciones es indicativa de resistencia a la terapia. Se apreciará que esta misma metodología puede usarse para determinar si el virus (y por tanto el paciente), será sensible o resistente a otros agentes anti-virales, tales como un inhibidor de proteasa, un inhibidor de la porción de serina-proteasa del gen NS3, y un inhibidor de la porción de helicasa del gen NS3.

En otra realización, la invención se refiere a un método de determinar la eficacia de un agente para tratar el HCV, comprendiendo el método comparar las secuencias de ácidos nucleicos contenidas en una primera muestra obtenida de un paciente expuesto al agente, en donde las secuencias de ácidos nucleicos son productos amplificados de una o más secuencias de cebadores de ácidos nucleicos de la presente invención, con las secuencias de ácidos nucleicos contenidas en una segunda muestra obtenida de un paciente que no ha estado expuesto a un agente, en donde las secuencias de ácidos nucleicos son productos amplificados de una o más secuencias de cebador de ácidos nucleicos de la presente invención, en donde un número aumentado de mutaciones en las secuencias de ácidos nucleicos de la primera muestra, con respecto a la segunda muestra, es una indicación de que el agente no es eficaz en tratar el HCV.

La práctica de la presente invención empleará, salvo que se indique otra cosa, técnicas convencionales de química, biología molecular, microbiología, DNA recombinante, e inmunología, que son conocidas por los expertos en la técnica. Dichas técnicas están explicadas completamente en la literatura científica. (Véase por ejemplo, Maniatis, Fitch & Sambrook, Molecular Cloning; A Laboratory Manual (1982); DNA Cloning, Volumes I and II (D. N Glover ed. 1985); Oligonucleotide Synthesis (M. J. Gait ed, 1984); Nucleic Acid Hybridization (B. D. Hames & S. J. Higgins eds. 1984); la serie Methods in Enzymology (Academic Press, Inc.), particularmente el Vol. 154 y el Vol. 155 (Wu and Grossman, and Wu, eds., respectivamente). Todas las patentes, solicitudes de patentes, y publicaciones mencionadas en la presente memoria, tanto en lo que antecede como en lo que sigue se incorporan en la presente memoria como referencias

Breve descripción de los dibujos

La Figura 1 representa los resultados de una PCR anidada de la proteasa NS3 del virus de la hepatitis C de 7 muestras de pacientes. El producto NS3 tiene 2746 pb. La escala de pesos moleculares del DNA está en la pista 1.

Las Figuras 2A-2B representa árboles filogenéticos construidos a partir de 30 cepas de proteasa de NS3 del HCV usando Análisis Parsimony y empalme de vecinos (Neighbor Joining). La Figura 2A representa un árbol de consenso del análisis Parsimony, mientras que la Figura 2B representa un árbol de remuestreo del análisis de empalme de vecinos.

Descripción detallada de la invención

El término "agente" incluye cualquier sustancia conocida adecuada para tratar una infección o enfermedad, en particular, agentes conocidos para tratar infecciones virales. El término "agente" está destinado además a cubrir agentes antivirales incluyendo, pero sin limitación, un inhibidor de proteasa, un inhibidor de la porción de de serina-proteasa del gen NS3, un inhibidor de la porción de helicasa del gen NS3, alfa-interferón, ribavirina e interferón pegilado.

Los términos "amplificación" o "amplificar" incluyen las reacciones necesarias para aumentar el número de copias de una secuencia de ácido nucleico (por ejemplo, una secuencia de DNA). Para los fines de esta invención, la amplificación se refiere al aumento exponencial in vitro del número de copias de una secuencia de un ácido nucleico diana, tal como el mediado por una reacción de amplificación por polimerasa, tal como por ejemplo, PCR, sin embargo, otras reacciones de amplificación abarcadas por la invención incluyen, por ejemplo, PCR con reacción

previa con transcriptasa inversa (abreviadamente RT-PCR por sus iniciales en inglés) (véase, por ejemplo, la patente de EE.UU. N° 4.683.202 de Mullis et al.), y la reacción en cadena de la ligasa (Barany, Proc. Natl. Acad. Sci. USA 88: 189-193 (1991)).

5 Las expresiones "Virus de la Hepatitis C" y "HCV" se refieren a la especie viral que es el principal agente etiológico de BB-NANBH, cuyo prototipo aislado se identifica en el documento WO89/046699; la publicación de la EPO 318.216; la patente de EE.UU. N° 5.350.671, cuyas descripciones se incorporan como referencia en la presente memoria. "HCV" tal como se usa en el presente texto incluye las cepas patógenas capaces de causar hepatitis C, y las cepas atenuadas o defectuosas que interfieren con las partículas derivadas de las mismas. El genoma del HCV está constituido por RNA. Se sabe que los virus que contienen RNA tienen tasas relativamente altas de mutación espontánea, pues se han descrito del orden de 10^{-3} a 10^{-4} por nucleótido incorporado (Fields & Knipe, "Fundamental Virology" (1986, Raven Press, N. Y.)). Puesto que la heterogeneidad y fluidez del genotipo son características inherentes a los virus de RNA, habrá múltiples cepas/aislados, que pueden ser virulentos o no virulentos, dentro de las especies de HCV.

15 Los términos "hibridan" o "hibridación" son conocidos en la técnica e incluyen la unión por puentes de hidrógeno de secuencias de DNA y/o RNA complementarias para formar una molécula dúplex.

El término "kit" es cualquier artículo (por ejemplo, un paquete o envase) que comprende al menos un reactivo, por ejemplo, un cebador, para amplificar específicamente y/o secuenciar una porción del genoma del HCV, siendo dicho artículo promocionado para la venta, distribuido o vendido como una unidad para realizar los métodos de la presente invención. Un kit también puede incluir instrucciones para su uso.

20 El término "marcador" se refiere a un resto molecular capaz de ser detectado, incluyendo, a modo de ejemplo, sin limitación, marcadores radiactivos que pueden ser incorporados por métodos conocidos (por ejemplo, traslado de muescas), isótopos radiactivos, biotina, grupos fluorescentes, grupos quimioluminiscentes (por ejemplo, dioxetanos, particularmente dioxetanos activados), digoxigenina, enzimas, anticuerpos, agentes luminiscentes, agentes precipitantes, colorantes, y similares.

25 El término "nucleótidos" se refiere a cualquier nucleótido (incluyendo nucleótidos modificados, por ejemplo, nucleótidos metilados o biotinilados) que pueden ser incorporados en un ácido nucleico por una polimerasa.

30 Como se usa en la presente memoria "ácido nucleico" incluye moléculas de DNA (por ejemplo, cDNA o DNA genómico), moléculas de RNA (por ejemplo, mRNA), y análogos del DNA o RNA generadas usando análogos de nucleótidos o usando química de ácidos nucleicos. Las modificaciones químicas incluyen metilación, biotinilación, y otras modificaciones conocidas en la técnica. Además, la molécula de ácido nucleico puede ser monocatenaria o bicatenaria.

El término "ácido nucleico diana" o "molde" incluye cualquier ácido nucleico destinado a ser copiado en, por ejemplo, una reacción de amplificación por polimerasa, tal como PCR.

35 El término "sonda" se refiere a una estructura constituida por un polinucleótido que forma una estructura híbrida con una secuencia diana, debido a la complementariedad de al menos una secuencia de la sonda con una secuencia de la región de la diana. Las regiones de los polinucleótidos de las sondas pueden estar compuestas de DNA, y/o RNA, y/o análogos de nucleótidos sintéticos. Están incluidas dentro de las sondas la "sondas de captura", "las sondas de bloqueo", "las sondas marcadoras".

40 El término "cebador" o "cebador de ácidos nucleicos" o "secuencia de cebador de ácidos nucleicos" incluye oligonucleótidos monocatenarios que, típicamente, tienen entre aproximadamente 4 hasta aproximadamente 100 bases, o alternativamente entre aproximadamente 17 a 30 bases, o alternativamente 20 o más bases, y están diseñados para hibridarse con un ácido nucleico molde correspondiente. Las moléculas de cebador pueden ser complementarias para una cadena con sentido o anti-sentido de un ácido nucleico molde y típicamente se usan como pares complementarios que flanquean una región de ácido nucleico de interés.

45 El término "polimerasa incluye una cualquiera de, o una mezcla de, las enzimas que polimerizan nucleótidos: DNA polimerasa I de E. coli, el fragmento de Klenow de DNA-polimerasa I de E. coli, DNA-polimerasa de T4, transcriptasa inversa en donde el molde es RNA y el producto de extensión es DNA, o una DNA-polimerasa termoestable. La expresión "DNA-polimerasa termoestable" incluye DNA-polimerasa termoestable aislada de *Thermus aquaticus*, *Thermus thermophilus*, *Thermus filiformis*, *Thermus flavus*, *Pyrococcus furiosus*, *Thermococcus litoralis*, una especie de *Thermotoga* o una de sus formas recombinantes.

50 El término "muestra" o "muestra biológica" se refiere a una muestra de tejido o fluido aislada de un individuo, incluyendo pero sin limitación, sangre, plasma, suero, fluido de la médula espinal, fluido linfático, las secciones externas de la piel y de los tractos respiratorio, intestinal y genitourinario, lágrimas, saliva, leche, orina, células de la sangre, tumores, fluidos amnióticos, órganos, DNA, RNA, o cDNA genómicos en solución o unidos a un sustrato, y también muestras de los constituyentes de un cultivo de células in vitro incluyendo, pero sin limitación, el medio acondicionado que resulta del crecimiento de las células en el medio de cultivo, células que se supone que están

infectadas por virus, células recombinantes, y componentes celulares, por ejemplo, cromosoma(s), orgánulos, tejidos incrustados en parafina o membranas aisladas de una célula.

La "cadena con sentido" de un ácido nucleico contiene la secuencia que tiene homología de secuencia con la del mRNA. La "cadena antisentido" contiene una secuencia que es complementaria de la de la "cadena con sentido".

- 5 La expresión "región diana" se refiere a una región del ácido nucleico que ha de ser amplificada y/o detectada. La expresión "secuencia diana" se refiere a una secuencia con la cual una sonda o cebador formará un híbrido estable en condiciones deseadas.

10 La expresión "secuencia polinucleotídica para la formación de diana" como se usa en la presente memoria, se refiere a una secuencia polinucleotídica que está compuesta de nucleótidos que son complementarios a una secuencia de nucleótidos diana; la secuencia tiene una longitud y complementariedad suficientes con la secuencia diana para formar un dúplex que tiene suficiente estabilidad para el fin deseado.

15 El HCV tiene al menos seis genotipos, y múltiples subtipos dentro de cada genotipo. Las secuencias que corresponden al HCV subtipos 1a y 1b pueden seleccionarse de las SEC ID NO: 1-23 y la SEC ID NO: 64, descritas en las Tablas 1A y 1G. Las secuencias que corresponden al HCV subtipo 2a pueden seleccionarse de las SEC ID NO: 24-27, descritas en la Tabla 1B. Las secuencias que corresponden al HCV subtipo 2b pueden seleccionarse de las SEC ID NO: 28-42, descritas en la Tabla 1C. Las secuencias que corresponden al HCV subtipo 3a pueden seleccionarse de las SEC ID NO: 43-46, descritas en la Tabla 1D. Las secuencias que corresponden al HCV subtipo 3b pueden seleccionarse de las SEC ID NO: 47-50, descritas en la Tabla 1E. Las secuencias que corresponden al HCV subtipo 4a pueden seleccionarse de las SEC ID NO: 51-63, descritas en la Tabla 1F. Las secuencias de la invención, descritas como las SEC ID NO: 1-64, se definen en la Lista de secuencias de la solicitud.

20 En una realización, los oligonucleótidos de la presente invención son al menos 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o más idénticos a la secuencia de nucleótido descritas en la SEC ID NO: 3. En otra realización, la molécula de ácido nucleico de la SEC ID NO: 3 tiene una longitud de al menos 4, 8, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 o más nucleótidos. Se apreciará que el extremo 5' puede contener mayor variabilidad (por ejemplo, sustituciones de ácidos nucleicos), pero permanece funcional (por ejemplo, capaz de hibridarse a la porción de los genes NS3 y NS4 del genoma del HCV).

25 La preparación de la secuencia de ácido nucleico de la presente invención se realiza por medios conocidos en la técnica, incluyendo, por ejemplo, métodos que incluyen escisión, transcripción o síntesis química. Las secuencias y/o regiones dianas del genoma que se seleccionan para las cuales son complementarias las secuencias de ácidos nucleicos para formación de diana dependen del fin. Por ejemplo, si el objetivo es escrutar la presencia de HCV en muestras biológicas (por ejemplo, sangre), la secuencia de ácido nucleico preferida sería usada como cebador, y se hibridaría a las regiones conservadas del genoma del HCV. Algunas de las regiones conservadas del genoma del HCV a las cuales pueden unirse la secuencia de ácido nucleico de la presente invención se describen en la presente memoria. Otras regiones del genoma que están conservadas son deducibles por comparación con las secuencias de nucleótidos de diversos aislados de HCV como se describe en la presente memoria.

30 Se apreciará que la secuencia de ácido nucleico de la presente invención no necesita consistir solamente en la secuencia que sea complementarias a la secuencia del HCV diana. Por tanto, la secuencia de ácido nucleico de la presente invención puede contener además, secuencias de nucleótidos u otros restos que sean adecuados para los fines para los cuales se usan las secuencias de ácidos nucleicos. Por ejemplo, las secuencias de ácidos nucleicos usadas como cebadores puede ser usada para la amplificación de secuencias del HCV por PCR, y pueden contener secuencias que, cuando están en forma de dúplex, forman sitios para las enzimas de restricción que facilitan la clonación de las secuencias amplificadas.

35 La secuencia de ácido nucleico de la invención puede ser usada como sonda para detectar la presencia de polinucleótidos del HCV (por ejemplo en el escrutinio de sangre contaminada), la muestra biológica que se ha de analizar, tal como sangre o suero, puede ser tratada, si se desea, para extraer los ácidos nucleicos contenidos en ella. El ácido nucleico de la muestra puede ser sometido a electroforesis en gel u otras técnicas de separación por tamaños; alternativamente, la muestra de ácido nucleico puede ser sometida a transferencia de mancha sin separación por tamaños. Para formar dúplex híbridos con la secuencia de formación de diana de la sonda, la región dianizada del ácido nucleico debe ser de forma monocatenaria. Cuando la secuencia está presente de modo natural en la forma monocatenaria, no se requerirá la desnaturalización. Sin embargo, cuando la secuencia esté presente en forma bicatenaria, la secuencia será desnaturalizada. La desnaturalización puede llevarse a cabo por diversos métodos conocidos en la técnica. Después de la desnaturalización el ácido nucleico analito y la sonda se incuban en condiciones que favorecen la formación de híbridos estables de la secuencia diana de la sonda con la supuesta secuencia dianizada del analito, y se detectan los dúplex resultantes que contienen la(s) sonda(s).

40 La detección del dúplex resultante, si existe, se consigue usualmente por el uso de sondas marcadas; alternativamente, la sonda puede estar sin marcar, pero puede ser detectable por unión específica con un ligando que está marcado, directa o indirectamente. Los marcadores adecuados y los métodos para marcar sondas y

ligandos son conocidos en la técnica e incluyen, por ejemplo, marcadores radiactivos que pueden ser incorporados por métodos conocidos (por ejemplo, traslado de muesca o tratamiento con quinasas), isótopos radiactivos, biotina, grupos fluorescentes, grupos quimioluminiscentes (por ejemplo, dioxetanos, particularmente dioxetano activados), digoxigenina, enzimas, anticuerpos, agentes luminiscentes, agentes precipitantes, colorantes, y similares.

- 5 Las secuencias de ácidos nucleicos de HCV amplificadas generadas por los métodos descritos en la presente memoria pueden ser clonados en un vector hospedante con un promotor de expresión y usarse como una vacuna de DNA para amplificar una respuesta del paciente al HCV, proporcionando de este modo una vacuna de DNA específica del paciente para el HCV. En otra realización, el clon puede ser usado para generar RNA para uso como vacuna antisentido. En una realización, se proporciona una vacuna para el HCV y se emplea como un agente inmunoterapéutico para la prevención del HCV. En otra realización, se proporciona una vacuna para el HCV y se emplea como un agente inmunoterapéutico para el tratamiento del HCV.

15 A modo de ejemplo, una vacuna para el HCV puede emplearse para la prevención y/o el tratamiento de HCV en un sujeto administrando la vacuna por una variedad de vías, por ejemplo, intradérmicamente, subcutáneamente o intramuscularmente. Además, la vacuna para el HCV puede ser administrada junto con adyuvantes y/o inmunomoduladores para reforzar la actividad de la vacuna y la respuesta del sujeto. En una realización, los dispositivos y/o composiciones que contienen la vacuna, adecuada para liberación prolongada o intermitente podrían ser implantados en el cuerpo a aplicados tópicamente al mismo para la liberación relativamente lenta de dichos materiales en el cuerpo. La vacuna para el HCV puede ser administrada junto con compuestos inmunomoduladores, que pueden alterar el tipo de respuesta inmune producida para obtener una respuesta que será más eficaz en eliminar el virus.

EJEMPLOS ESPECÍFICOS

EJEMPLO 1

DIVERSIDAD GENÉTICA Y RESPUESTA AL INTERFERÓN DEL GEN DE LA PROTEASA NS3 DE LAS CEPAS CLÍNICAS DEL VIRUS DE LA HEPATITIS C.

- 25 Este ejemplo describe un análisis de la diversidad genética natural dentro del gen NS3, en donde se desarrolló un método de PCR anidado para obtener datos de la secuencia del gen NS3 del HCV directamente de cepas de pacientes. Los datos se analizaron para determinar la diversidad genética, la filogenia de los virus, y la selección de variantes genéticas por terapia con interferón con o sin ribavirina. Se determinaron los efectos de la diversidad genética sobre la estructura de la enzima usando modelación molecular. Por tanto, se desarrolló un método
- 30 comprensivo para facilitar el análisis de la variación genética natural dentro de este gen y sus efectos sobre la estructura de las proteínas.

La tercera parte del extremo N-terminal del gen 3 no estructural (NS3) del virus de la hepatitis C codifica una serina-proteasa (véase la Figura 1 en Holland-Staley, C. A., 2002, supra).

- 35 Aunque la diversidad genética ha demostrado ser extensa en otras regiones del genoma del HCV, tal como la región de la envoltente, la extensión de la diversidad dentro de la región de la proteasa NS3 del HCV es ampliamente desconocida. Por consiguiente, los datos expuestos en la presente memoria se usaron para determinar las tasas evolutivas del virus y para identificar los efectos de las mutaciones sobre la estructura de las enzimas. Para investigar la diversidad genética natural de esta enzima se desarrolló una reacción PCR anidada para obtener los datos de la secuencia de la proteasa NS3 directamente de cepas de pacientes. Estos datos se usaron para
- 40 determinar la diversidad genética, las tasas filogenéticas y evolutivas, y la selección de variantes por terapia con interferón. También se intentó averiguar el efecto potencial de la diversidad genética sobre la estructura de la enzima usando modelación molecular. Tras realizar estos experimentos, se ha establecido una base de datos de secuencias naturales. Esta base de datos es útil en el desarrollo de una nueva terapia antiviral.

- 45 La diversidad genética entre las cepas es importante porque pone de manifiesto que algunos tipos de HCV son más patógenos que otros, Cooreman, M. P. et al., (1996) Hepatitis C virus: biological and clinical consequences of genetic heterogeneity. *Scand. J. Gastroenterol. Suppl* 218: 106-115; Dusheiko, G. et al., (1994) Hepatitis C virus genotypes: an investigation of type-specific differences in geographic origin and disease. *Hepatology* 19: 13-18; Stempniak, M. et al., (1997). The NS3 proteinase domain of hepatitis C virus is a zinc-containing enzyme. *J. Virol.* 71: 2881-2886), es decir, el HCV tipo 1 es más patógeno que los otros tipos, y el subtipo 1b es más patógeno que el subtipo 1a. La causa de la mayor patogénesis de estas cepas no está clara. Describir las diferencias entre cepas que tengan un rápido progreso a la enfermedad en comparación con las que no lo tienen ayudará a elucidar la posible causa y los mecanismos de la patogénesis (Farci, P. et al., (2000) The outcome of acute hepatitis C predicted by the evolution of the viral quasispecies. *Science* 288: 339-344). Recientemente, Yanagi et al., (Yanagi, et al., (1998) Transcripts of a chimeric cDNA clonof hepatitis C virus genotype 1b are infectious in vivo. *Virology* 244: 161-172) mostraron que el número de mutaciones dentro del genoma del HCV puede estar directamente relacionado con la infectividad. Dichos autores eligieron tres clones de cDNA quiméricos de la misma fuente y los inocularon en chimpancés. De los tres clones sólo uno era infeccioso. Este clon contenía 3 sustituciones de aminoácidos cuando

se comparó con la cepa parental (precursora) mientras que los clones no infecciosos contenían 7 y 9 sustituciones de aminoácidos. Esto puede indicar una desviación de la selección en chimpancés debido a diferencias del sistema inmune en lugar de una característica de la cepa del HCV (Bassett, S. E. et al., (1998) Analysis of hepatitis C virus-inoculated chimpanzees reveals unexpected clinical profiles. *J. Virol.* 72: 2589-2599). En otro estudio, cuando se examinó la región 1 hipervariable (Shindo, M. et al., T (1996) The clinical significance of changes in genetic heterogeneity of the hypervariable region 1 in chronic hepatitis C with interferon therapy. *Hepatology* 24: 1018-1023) se mostró que las mayores cantidades de heterogeneidad estaban asociadas con una respuesta global más pobre a pesar del tratamiento con alfa-interferón. Además, Chambers, T. J. et al., (Chambers, T. J. et al. (1991) Processing of the yellow fever virus nonstructural polyprotein: a catalytically active NS3 proteinase domain and NS2B are required for cleavages at dibasic sites. *J. Virol.* 65: 6042-6050); establecieron que las mutaciones que afectan a la actividad de la proteasa NS3 del virus de la fiebre amarilla, eran letales para la replicación viral. La conclusión de estos estudios afirma la necesidad de caracterizar más la función que desempeña la heterogeneidad en la infección por HCV. Las composiciones y métodos de la presente invención proporcionan un mecanismo para dicha caracterización. Filocamo et al., 1999. Selection of functional variants of the NS3-NS4A protease of hepatitis C virus (SBV) by using chimeric sindbis viruses. *J. Virol.* 73: 561-575 usaron virus Sindbis quiméricos para analizar las variantes de la proteasa NS3/NS4A, sin embargo, su estudio una proteasa clonada y se basaba en la maquinaria de replicación del SBV para producir variación de secuencias. Aunque esto permitió la caracterización de nuevas variantes de proteasas, estas variantes pueden o no pueden representar las obtenidas de la replicación del HCV in vivo. Además, Forns et al., (Forns X. , et al., (1997), How *Escherichia coli* can bias the results of molecular cloning: preferential selection of defective genomes of hepatitis C virus during the cloning procedure. *Proc. Natl. Acad. Sci. USA* 94: 13909-13914) encontraron cuando clonaban productos por PCR de proteínas estructurales del HCV en *E. coli*, que ocurría una fuerte selección para los clones defectuosos y no ocurría una representación precisa de la verdadera diversidad genética. La razón para esto no está clara; sin embargo, la toxicidad para el hospedante del *E. coli* por proteínas virales puede desempeñar un papel. La secuenciación directa de muestras de pacientes debe dar una representación más precisa de la verdadera diversidad genética.

La mayoría de los virus de RNA tiene una alta tasa de mutación debido a la falta de enzimas correctoras del texto leído y debido a polimerasas de RNA dependientes de RNA se piensa que producen un error aleatorio por ronda de replicación genómica (Drake, J. W. (1993) Rates of spontaneous mutation among RNA viruses. *Proc. Natl. Acad. Sci. USA* 90: 4171-4175). Por tanto, si hay 10⁷-8 virus producidos por día y la tasa de mutación es aleatoria, entonces cada sitio contendrá teóricamente un número igual de mutaciones. Los datos expuestos en la presente memoria, sin embargo, mostraron que no sucedía esto, indicando que aunque las mutaciones fueran aleatorias, algunas regiones del virus son más permisivas para las mutaciones que otras regiones. Esto sugiere que el virus, o al menos la región de la proteasa NS3, está sometido a otras restricciones funcionales. La localización de los cambios de aminoácidos en la proteasa NS3 del HCV de cepas clínicas indica que las sustituciones de aminoácidos tienen lugar en ambos dominios el N-terminal y el C-terminal. Ya se ha demostrado que el dominio N-terminal de NS3 exhibe una plasticidad estructural por la unión de NS4A. Kim, J. L. et al., (Kim, J. L. et al., (1996) Crystal structure of the hepatitis C virus NS3 protease domain complexed with a synthetic NS4A cofactor peptide. *Cell* 87: 343-355) y; Yan, Y. et al., (Yan, Y. et al., (1998) Complex of NS3 protease and NS4A peptide of BK strain hepatitis C virus: a 2.2 Å resolution structure in a hexagonal crystal form. *Protein Sci.* 7: 837-847) describieron que el dominio N-terminal es flexible puesto que la unión de NS4A conduce a la ordenación de los 28 residuos N-terminales y causa transposiciones locales para una conformación catalíticamente más favorable del sitio activo. La activación de la proteasa NS3 por el péptido NS4A conduce a una mejora de ~950 veces la hidrólisis catalítica de los péptidos que imitan el empalme NS4A/NS4B (Yan, Y. et al., 1998 supra). Basándose en los cambios estructurales predichos por las secuencias de las cepas clínicas del, se cree que habrá cambios funcionales con respecto a la proteasa de tipo silvestre. Un ejemplo de dicho cambio sería la alteración de la constante de velocidad de pseudo-segundo orden (k_{Cat}/KM) para las diversas cepas clínicas. Una segunda observación es que la variación estructural del complejo NS3/NS4A del HCV debe ser un factor en el diseño basado en la estructura de los inhibidores con el fin de retener la eficacia contra una amplia gama de aislados clínicos.

Para las cepas de los pacientes 23 y 25 las diferencias en la denominación del subtipo entre la región 5' no traducida y la región de la proteasa NS3 son desconocidas, pero pueden ocurrir debido a la presencia de más un subtipo. La infección simultánea con múltiples especies del virus del Dengue ha mostrado que ocurren (Gubler, D. J. et al., (1985) A case of natural concurrent human infection with two dengue viruses. *Am. J. Trop. Med. Hyg.* 34: 170-173). Para la infección con HCV esto es más probable debido a múltiples transfusiones de sangre o infecciones a través de abuso de drogas por vía intravenosa. Debido a que los subtipos del HCV 1a y 1b son los más predominantes en esta distribución geográfica, aumenta la probabilidad de que podrían recombinarse. Las pruebas de la recombinación se han mostrado en el virus del Dengue, en donde la diversidad genética ocurre cuando la RNA-polimerasa cambia entre dos diferentes genomas durante la replicación, dando como resultado un RNA híbrido (Holmes, E. C. et al., (2000) The causes and consequences of genetic variation in dengue virus. *Trends Microbiol.* 8: 74-77). Este puede ser el caso de las cepas 23 y 25, en donde son clasificadas como HCV tipo 1a basándose la región 5' no traducida, pero más próximas al clado (cada una de las ramas del árbol filogenético propuesto para agrupar a los seres vivos, es decir un conjunto de especies emparentadas con un antepasado común) HCV 1b por la secuencia de ácido nucleico de la NS3. Los resultados de la secuencia fueron netos y no mostraron pruebas de más de una especie, sugiriendo que la recombinación ocurrió antes del comienzo de NS3. La comparación de las

secuencias de aminoácidos, sin embargo, mostró resultados muy diferentes. El paciente 23 tenía 3 mutaciones cuando se comparó con la secuencia de consenso tipo 1a y 16 mutaciones cuando se comparó con el tipo 1b. El paciente 25 mostró resultados similares con 4 mutaciones cuando se comparó con el tipo 1a y 16 cuando se comparó con el tipo 1b. Se desconocen las diferencias entre los análisis de las secuencias de nucleótidos y el análisis de proteínas. Sin estar limitado por la teoría, se cree que estas dos cepas representan un nuevo subtipo, que en el árbol evolutivo está situado entre los subtipos 1a y 1b.

Este ejemplo reveló que hay una variabilidad significativa en las cepas clínicas del HCV tanto en las secuencias de nucleótidos (30,2% para el subtipo 1a y 25,8% para el subtipo 1b) como en las de secuencias aminoácidos (12,2% para el subtipo 1a y 12,2% para el subtipo 1b). El análisis filogenético mostró dos distintos clados con dos aislados del HCV que se agrupan como un clado hermano en 1b. Además, el análisis estructural reveló que la mayoría de las mutaciones se encuentra en el N-terminal de la enzima. Cuando las cepas fueron clasificadas como si el paciente hubiera recibido o no recibido terapia antiviral, no se encontró diferencia en el número o localizaciones de las mutaciones en las cepas 1a. Sin embargo, las cepas 1b demostraron una caída global en el número de posiciones que estaban mutadas. Por tanto, este estudio demostró que hay diferencias significativas entre cepas naturales que pueden plantear un problema para el desarrollo de fármacos en base a la estructura, para el cual esta invención proporciona una solución.

Las relaciones entre variación de secuencias y cambios estructurales en el complejo NS3/NS4A del HCV también pueden detectarse por cristalografía de rayos X. Basándose en las composiciones y métodos de la presente invención, pueden diseñarse fármacos antivirales, incluyendo anti-HCV, y desarrollarse con el conocimiento de la variabilidad en la estructura del complejo NS3/NS4A a partir de las cepas clínicas del HCV para desarrollar finalmente terapias eficaces contra el HCV.

MATERIALES Y MÉTODOS

Extracción del RNA viral.

Se aisló RNA del HCV a partir de suero de pacientes usando el kit de mini-aislamiento de RNA viral QIAamp (de Qiagen Inc., Valencia, CA.). El aislamiento se realizó siguiendo las instrucciones de los fabricantes.

RT-PCR y amplificación de RNA viral.

Se usó RT-PCR seguida por PCR anidada de 2ª ronda para amplificar el gen NS3 completo del HCV subtipos 1a o 1b, a partir de cepas clínicas. Para la etapa de RT-PCR se usó el kit de RT-PCR Promega Access (de Promega Corporation, Madison, WI). Para esto, se diseñaron dos cebadores oligonucleotídicos que flanquean el gen NS3 del HCV. El cebador para 5' como se describe en la SEC ID NO: 1 se hibrida 881 pb antes del comienzo del gen NS3. El cebador para 3' descrito en la SEC ID NO: 2 se hibrida 338 pb aguas abajo (en dirección 3') del gen NS3 y se usa para iniciar la etapa de la transcriptasa inversa (RT), (Tabla 1G, SEC ID NO: 1-4, SEC ID NO: 64 y SEC ID NO: 6, sub-apartado "amplificación"). Esto permite la producción de cDNA seguido por la amplificación inicial de la región deseada. La mezcla de amplificación que contenía 25 pmol de cada cebador, diversos dNTP (desoxitriofosfatos de nucleótidos) 200µM, 2,5 U de RT-polimerasa del virus del mosaico de la alfalfa (AMV), 2,5 U de DNA-polimerasa de Tfl, MgSO4 1,5 mM, y 20 µl de RNA viral se añadió a un aparato de ciclo térmico pre-calentado (48°C) modelo Perkin-Elmer 9700 (de Perkin Elmer Cetus, Foster City, CA). El protocolo de la PCR consistió en una etapa a 48°C durante 45 minutos, seguido por una desnaturalización inicial a 94°C durante 2 minutos, y 35 ciclos de 94°C durante 15 segundos; 55°C durante 20 segundos; 72°C durante 2 minutos; y una extensión final a 72°C durante 10 minutos.

Para la segunda etapa de la PCR, se amplificó el producto de la PCR de la 1ª ronda usando cebadores que se hibridan dentro de la reacción previa, creando una amplificación 'anidada'. La reacción PCR de la segunda ronda usó el cebador para 5' descrito como SEC ID NO: 3 que se hibrida 693 pb aguas arriba (en dirección 5') de NS3 y el cebador 3' descrito como SEC ID NO: 4 que se hibrida 161 pb aguas abajo (en dirección 3'). La mezcla de amplificación que contenía 25 pmol de cada cebador, diversos dNTP 200 µM, 2,5 U de DNA-polimerasa Taq®, MgSO4 1,5 mM, y 5 µl del producto de la 1ª ronda se añadió a un aparato de ciclo térmico precalentado (94°C). La desnaturalización inicial consistía en 10 minutos a 94°C, seguido por 35 ciclos de 94°C durante 15 segundos; 55°C durante 20 segundos; 72°C durante 2 minutos y una extensión final a 72°C durante 10 minutos. El producto de amplificación resultante es una sola banda de 2746 pb sobre un gel de agarosa al 1%. El gen NS3 completo se amplificó usando este método, sin embargo, solamente se caracteriza en la presente invención el gen de la proteasa NS3. Los productos de la segunda etapa de reacción se purificaron y concentraron usando el Gene Clean Spin Protocol (de Bio 101, Vista, CA). Los productos purificados se usaron directamente para secuenciar el DNA. Para asegurarse de que cada aislamiento del RNA del HCV era satisfactorio, se usaron como control cebadores que cubren la región 5'-no traducida altamente conservada (5'-UTR) (Yuki, N. et al., (1997) Hepatitis C virus replicative levels and efficiency of genotyping by specific PCR and antibody assay. J. Clin. Microbiol. 35: 1184-1189). Los productos de PCR de esta región de control también se purificaron como se describe antes y se secuenciaron. Los datos de las secuencias de la región 5'UTR se usaron luego para determinar el genotipo de cada cepa del HCV (O'Brien, C. B. et al., D (1997) cDNA sequencing of the 5'-noncoding region (5'NCR) to determine hepatitis C genotypes in patients with chronic hepatitis C. Digestive Diseases and Sciences 42: 1087-1093). Si la segunda

ronda de la amplificación por PCR fuera no satisfactoria, se realizó una nueva ronda de amplificación usando un segundo conjunto de cebadores de seguridad ('backup') internos. El cebador de seguridad 5' descrito como SEC ID NO: 64 y el cebador de seguridad 3' descrito como SEC ID NO: 6 dan como resultado un producto de amplificación de 2377 pb, que corta el gen NS3 (región de helicasa) acortándolo en 19 pb.

5 Secuenciación de DNA

La secuenciación de los productos de reacción purificados de la segunda etapa de la PCR usando la tecnología ABI Big Dye Terminator (de Applied Biosystems, Foster City, CA). Se usaron cinco cebadores como los descritos en las SEC ID NO: 7-11 (véase la Tabla 1G, sub-apartado "Secuenciación") diseñados para cubrir la región de la proteasa NS3 en ambas cadenas con sentido y anti-sentido. Para el análisis con ABI, las reacciones de secuenciación se llevaron a cabo en placas de microtitulación usando un aparato de ciclo térmico durante 25 ciclos de 96°C durante 10 segundos; 50°C durante 5 segundos; 60°C durante 4 minutos. Los productos de reacción se purificaron con NaOAc/EtOH, se volvieron a suspender en tampón de carga, se desnaturalizaron y se secuenciaron en un secuenciador ABI 377 (de Applied Biosystems).

Análisis de secuencias.

15 Los resultados de las secuencias obtenidos por el secuenciador ABI se analizaron usando el programa informático de análisis de secuencias ABI versión 3.3. Las secuencias individuales se alinearon frente a una secuencia de consenso derivada de datos de las secuencias de todos los pacientes y frente a los tipos de HCV 1a y 1b (Números de acceso AF009606 y AJ000009) (Yanagi, M. et al., (1997) Transcripts from a single full-length cDNA clone of hepatitis C virus are infectious when directly transfected into the liver of a chimpanzee. Proc. Natl. Acad. Sci. USA 94: 8738-8743; and Yanagi, M. (1998). Transcripts of a chimeric cDNA clone of hepatitis C virus genotype 1b are infectious in vivo. Virology 244: 161-172). Para comparación se eligieron las dos secuencias de Genbank porque estaban disponibles como construcciones clonadas para uso como controles positivos para las subsiguientes valoraciones de la expresión de proteínas (no se muestran los datos). Se comparan los aminoácidos que difieren de la secuencia de consenso y se usan para determinar que residuos se conservan y cuales son variables.

25 Análisis filogenético

Se alinearon las secuencias usando el programa Sequencer 3.0, y el alineamiento múltiple se exportó a un archivo Nexus. El análisis Parsimony se llevó a cabo usando Phylogenetics Analysis Using Parsimony (PAUP) 3.1. 1 (Swofford, D. L. (1993) PA UP: Phylogenetic analysis using Parsimony. 3.1 ed. Illinois Natural History Survey, Champaign, IL) con una búsqueda heurística con permutación de ramas. Para comparación, también se usó una rama y opción vinculante para buscar los árboles más parsimoniosos. Se generó una matriz de distancia basada en las distancias de dos parámetros de Kimura usando el programa DNADIST de PHYLIP Versión 3.573c (Felsenstein, J. (1993) PHYLIP (Phylogenetic Inference Package), 3.5p ed. Department of Genetics, University of Washington, Seattle). Se usó una relación de transición a transversión dos a uno. La matriz de distancia se analizó luego por análisis de empalme de vecinos [Neighbor-Joining Analysis (NJ)] usando el programa NEIGHBOR en PHYLIP. El análisis de soporte con remuestreo se ejecutó usando 1000 re-muestréos para ambos análisis NJ y Parsimony. Los conjuntos de datos múltiples para el análisis NJ se generaron por el programa SEQBOOT y los árboles múltiples se determinaron por consenso mediante el programa CONSENSE, ambos en PHYLIP. El remuestreo de Parsimony se completó usando la opción de remuestreo en PAUP.

40 Se contó el número de sitios de segregación para cada población a partir de los alineamientos múltiples generados por Sequencer 3.0. La distinción de sustituciones sinónimas frente a sustituciones no sinónimas se hizo basándose en las traducciones propuestas de las secuencias. Se estimó el parámetro de mutación de la población μ de acuerdo con Watterson (Watterson, G. A. (1975) On the number of segregating sites in genetical models without recombination. Theoretical Pop. Biol. 7: 256-276). Tests of neutrality follow McDonald and Kreitman (McDonald, J. et al., (1991) Adaptive protein evolution at adh locus in Drosophila. Nature 351: 652-654).

45 Modelización de la estructura de las proteínas por homología

La modelización de la estructura de las proteínas por homología mediante satisfacción de las restricciones espaciales se realizó de acuerdo con el método de Sali, 1995 (Sali, A. (1995) Comparative protein modeling by satisfaction of spatial restraints. Mol. Med. Today 1: 270-277). La modelización de la estructura de las proteínas por homología se basa en la construcción de un modelo estructural de una proteína sobre la base de la estrecha similitud con una proteína molde de estructura conocida.

55 En la primera etapa de la modelización de la estructura de una proteína se obtuvieron los alineamientos entre la secuencia desconocida y las estructuras de molde relacionadas. En segundo lugar se derivaron las restricciones de diversas distancias, ángulos y ángulos diédricos en la secuencia basándose en su alineamiento con las estructuras del molde. Finalmente, se obtuvieron los modelos tridimensionales minimizando las violaciones de restricciones de energía y derivadas de la homología, usando gradientes conjugados de métodos de dinámica molecular. Una etapa importante en los experimentos de modelización de proteínas por homología es la evaluación de la calidad del modelo. Para asegurar la calidad se usaron comprobaciones de la consistencia interna como las implementadas en

el paquete de programas informáticos MODELLER 4 (Sali, A. et al., (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* 234: 779-815) para analizar el perfil tridimensional.

5 Se usó el paquete de programas informáticos para la modelización comparativa de proteínas MODELLER para calcular las estructuras del complejo NS3/NS4 del HCV. Este paquete de programas informáticos ha sido analizado exhaustivamente en proyectos genómicos estructurales por Sánchez and Sali, 1998 (Sánchez, R. et al., (1998) Large-scale protein structure modeling of the *Saccharomyces cerevisiae* genome. *Proc. Natl. Acad. Sci. USA* 95: 13597-13602). La estructura cristalina de 2,2Å del complejo de la proteasa NS3 del HCV y el péptido NS4 (Yan, Y. et al., 1998, supra) se usó como molde para los experimentos de modelización por homología para examinar los cambios estructurales en las cepas clínicas de la proteasa NS3 del HCV. Las estructura cristalina del complejo 10 NS3/NS4A del HCV y las estructuras modeladas de las proteasas NS3/NS4A del HCV se superpusieron por mínimos cuadrados usando la suite de programas O (Jones, T. A. et al., (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Cryst.* A47: 110-11932).

15 La optimización de los modelos de NS3/NS4A del HCV se llevó a cabo usando ajustes por defecto en el programa informático MODELLER 4.0 (número de iteraciones en la optimización = 200; tipo de limitaciones no vinculantes = términos de repulsión de esferas blandas dinámicas). El modelo se obtiene optimizando una función objetiva (limitaciones espaciales combinadas y términos de energía CHARMM que impone la estereoquímica adecuada) en el espacio cartesiano.

Números de acceso a GenBank de las secuencias de nucleótidos.

20 Las secuencias de nucleótidos descritas en la presente memoria han sido depositadas en la American Type Culture Collection (ATCC), 10801 University Boulevard, Manassas, VA 20110-2209, el 9 de abril de 2001 y se les ha asignado los números de acceso AF369214 a AF369263. Estos depósitos se mantendrán bajo los términos del Tratado de Budapest referente al reconocimiento internacional del depósito de microorganismos para fines del procedimiento de patentes. Estos depósitos se hicieron simplemente para comodidad de los expertos en la técnica y no se reconoce que se requiera un depósito bajo las normas del 35 U. S. C. §112.

25 RESULTADOS

Población de pacientes

30 Se analizó suero de treinta pacientes infectados con HCV subtipos 1a ó 1b. El grupo incluía dieciocho varones y doce mujeres de una edad media de 50,2 años (intervalo 21-76). Diecinueve no tuvieron tratamiento, mientras que los once restantes 11 tuvieron tratamiento con α -IFN solo (pacientes 11, 183 y 205), α -IFN + ribavirina (paciente 176) o una ronda inicial de α -IFN seguida por una segunda ronda de terapia de combinación (pacientes 252, 174, 186, 1D, 1Y, 179, 25). Todos los pacientes recibieron 3 millones de unidades, 3 veces a la semana durante 12 meses excepto los pacientes 11 y 205, que recibieron un segundo curso de terapia con α -IFN de 5 millones de unidades 4 veces a la semana y 1,5 millones de unidades 3 veces a la semana, respectivamente, y los pacientes 186 y 1Y, a los que se les interrumpió la terapia después de 6 meses. Ninguno tuvo una respuesta prolongada. Todos los pacientes 35 tuvieron cargas virales de RNA del HCV de entre 5,1 x 10⁴ y 2,3 x 10⁶ copias/ml determinadas usando el Roche Amplicor® HCV Monitor Test.

Determinación del genotipo.

40 La región 5' no traducida y conservada se secuenció para clasificar cada cepa en el genotipo respectivo usando el método de O'Brien et al., (O'Brien, C. B. et al., (1997) cDNA sequencing of the 5'-noncoding region (5'NCR) to determine hepatitis C genotypes in patients with chronic hepatitis C. *Digestive Diseases and Sciences* 42: 1087-1093). Los resultados del genotipo mostraron que 21 secuencias eran HCV subtipo 1a y 9 eran del subtipo 1b. Estos datos son consistentes con la población de pacientes, siendo los subtipos 1a y 1b los más prevalentes en esta región del mundo. Los datos expuestos en la presente memoria demuestran una correlación entre el genotipo 5'-UTR y los datos de la secuencia de NS3.

45 Amplificación y secuenciación de la región de NS3 de cepas clínicas.

50 El gen de la proteasa NS3 ha sido secuenciado satisfactoriamente y amplificado en la presente invención a partir de los sueros de los 30 pacientes infectados con HCV subtipos 1a o 1b. El producto de la amplificación produjo una banda neta con una tasa de éxito de >80% (Figura 1). Luego se tradujeron las secuencias resultantes y se derivó una secuencia de consenso por alineamiento de las secuencias de todos los pacientes. Cada secuencia de paciente se comparó con la secuencia de consenso o frente a las secuencias prototipo encontradas en Genbank para cada subtipo (Nº AF009606 y Nº AJ000009 de Genbank) (Yanagi, M. 1998, supra; véase también Yanagi, M. et al., 1998, supra). Las regiones variables y conservadas dentro de estas secuencias se determinaron luego tanto para los residuos de nucleótidos como de aminoácidos.

Análisis de las secuencias de nucleótidos.

Hay 543 nucleótidos que codifican el gen de la proteasa NS3. Los datos de la secuencia se analizaron para determinar la distribución de nucleótidos en cada posición cuando se comparó con la secuencia de consenso derivada de los pacientes y las dos secuencias prototipo de Genbank. Para aquellas cepas que pertenecían al HCV subtipo 1a, 164 posiciones contenían una o más sustituciones de nucleótidos cuando se compararon con la secuencia de consenso de los pacientes, dando como resultado una variación de secuencia global de 30,2% (datos no mostrados). Estas incluían 133 transiciones, 9 transversiones y 22 sitios con ambas, transiciones y transversiones. Cuando se compararon con el prototipo 1a, 20 posiciones eran diferentes (variación de secuencia global de 3,5%), con 17 posiciones que contenían transiciones y 3 que contenían transversiones (Tabla 2A). No se detectaron inserciones o deleciones. Para cepas clasificadas como HCV subtipo 1b, 140 posiciones contenían una o más sustituciones cuando se compararon con la secuencia de consenso de los pacientes, para una variación de secuencia global de 25,8%. Había 114 transiciones, 8 transversiones y 18 sitios que contenían ambas, transiciones y transversiones. Había 38 posiciones que variaban desde la secuencia 1b prototipo (global 7,4%), con 34 transiciones y 4 transversiones (Tabla 2B). En las secuencias no se detectaron inserciones o deleciones. Sin vincularse a la teoría, debido a que la tasa de mutación predicha introducida por la amplificación por PCR fue ~0,1%, se cree los datos reflejan una representación precisa de la variación de la secuencia genética natural.

Análisis de las secuencias de aminoácidos.

Hay 181 aminoácidos que comprenden la proteasa NS3. Los efectos de la sustitución de nucleótidos sobre la secuencia de aminoácidos mostraron que la mayoría de las sustituciones eran silenciosas. Para las secuencias de 21 pacientes que pertenecen al HCV subtipo 1a, 22 residuos o 12,2% diferían de la secuencia de consenso de los pacientes (Tabla 3). La secuencia prototipo difería de la secuencia de consenso de los pacientes en 5 residuos (A40T, K80Q, A91S, I153L, E176G) incluyendo una mutación no encontrada en ninguna secuencia de pacientes (subrayada). Tres sitios tenían más de dos especies de aminoácidos presentes (Vizmanos, J. L. 1998, supra; Yanagi, M. et al., (1997) Transcripts from a single full-length cDNA clon of hepatitis C virus are infectious when directly transfected into the liver of a chimpanzee. Proc. Natl. Acad. Sci. USA 94: 8738-8743), y un sitio (residuo 86) pudo tolerar 4 residuos diferentes. Incluyendo la secuencia de control, 9 sitios tenían sustituciones del mismo tipo de aminoácidos, 9 sitios pudieron acomodar aminoácidos neutros o polares, 3 sitios básicos o polares, 2 sitios ácidos o polares, y un sitio ácido o neutro. Solamente dos cepas fueron idénticas a la secuencia de consenso sin sustituciones de aminoácidos (K y 161).

Para cepas de pacientes que estaban clasificadas como HCV 1b, había una variación de secuencia global de 12,2%, con 22 sitios diferentes de la secuencia de consenso derivada de los pacientes (Tabla 4). La secuencia de consenso de los pacientes difería de la secuencia prototipo en 8 residuos (L13V, V48I, Q80L, L94M, T98S, A150V, I153V, V170I) incluyendo 4 sitios no identificados previamente (subrayados), ascendiendo el número total de sitios que contenían uno o más aminoácidos a 26. Más de dos especies de aminoácidos estaban presentes en 2 sitios (61 y 80), y el residuo 67 tenía hasta 4 residuos diferentes (el residuo 86 no contenía mutaciones). El mismo tipo de sustitución se vio en once sitios, 8 sitios acomodaban aminoácidos neutros o polares, 1 sitio básico o polar, 2 sitios básicos o neutros y el sitio restante tenía un cambio de aminoácidos de neutro a aromático. Solamente la cepa 153 era idéntica a la de consenso. Para los tipos 1a ó 1b no se encontraron sustituciones de aminoácidos en los residuos conservados conocidos.

Selección por terapia con interferón.

Las diferencias más significativas se apreciaron cuando las cepas se agruparon según si el hospedante había recibido o no había recibido terapia con interferón. Para las cepas que no habían sido sometidas a terapia con interferón con o sin ribavirina, el número global medio de mutaciones por paciente era 1,9 para las cepas del tipo 1a, con 14 sitios que contenían mutaciones. Después de la terapia con interferón, el número medio de mutaciones por paciente era 2,6 con 15 sitios que contenían diferentes aminoácidos. Solamente 7 sitios eran similares tanto antes como después de la terapia (Tabla 3). Para las cepas del tipo 1b, el número medio de mutaciones por paciente no sometidas a terapia con interferón era 3,6 con variación global en 18 sitios. Sin embargo, después de la terapia con interferón, aunque el número de mutaciones por paciente no fue significativamente diferente (3,0), el número de sitios que contenían mutaciones en las cepas del tipo 1b descendió a 10, o en 56% (Tabla 4) con 6 sitios similares en las cepas con y sin terapia. Esto sugiere que las cepas sensibles fueron eliminadas durante la terapia con interferón y estos 10 sitios restantes pueden de algún modo estar implicados en la resistencia del virus al interferón.

En resumen, los cambios de aminoácidos que ocurren más frecuentemente están representados por mutaciones de aminoácidos hidrófobos a hidrófobos seguidos por cambios de cadenas laterales de aminoácidos cargadas positivas a neutras. La tasa de mutaciones desde cadenas laterales de aminoácidos más voluminosas a menos voluminosas ocurre a una tasa superior a la inversa. En las cepas del tipo 1b, la diversidad de secuencias en la proteasa NS3 es significativamente reducida en cepas de pacientes que han experimentado terapia con IFN. Esto sugiere que las cepas sensibles son eliminadas durante la terapia con interferón y aquellas cepas que permanecen tras la terapia pueden ser menos sensibles al fármaco.

Análisis filogenético y evolución molecular.

El análisis Parsimony que usa una búsqueda heurística con permutación de ramas detectó 141 árboles igualmente parsimoniosos con 696 etapas. Una rama comparable y una opción vinculante detectaron 144 árboles igualmente parsimoniosos con 696 etapas. Las secuencias se agruparon en más de 95% de los árboles en dos clados, los clados HCV 1a y HCV 1b. El clado HCV 1a consiste en 19 cepas (K, 4, 11, 12, 24, 161, 170, 174, 176, 177, 183, 186, 194, 252, 1C, 1D, 1H, 1X, y 1Y) en lugar de 21 como las identificadas con la determinación del genotipo. Las muestras restantes (1, 23, 25, 153, 179, 205, 1A, 1G, 1U, 2D y 2F) estaban agrupadas en el clado HCV 1b. Estos resultados están sólidamente basados en análisis de remuestreo, aunque las relaciones jerárquicas dentro de los dos clados no están sólidamente basadas (Figura 2A). En particular, la Figura 2A representa un árbol de consenso del análisis Parsimony. Las longitudes de las ramas se dibujan para igualar los grupos jerárquicos y no reflejan las etapas mutacionales entre las cepas. Dentro del clado HCV 1b, las dos cepas 23 y 25 son taxones hermanos con respecto a las cepas restantes. Estas dos muestras, 23 y 25, se agrupan con el tipo HCV 1a por determinación del genotipo y comparación de las secuencias de aminoácidos.

La Figura 2B representa un árbol de remuestreo del análisis de empalme de vecinos. El análisis de empalme de vecinos produjo resultados muy similares a los del análisis Parsimony que indica un clado HCV 1b muy separado consistente en las cepas 1, 1G, 153, 2D, 1A, 179, 2F, 1U y 205 (Figura 2B). El soporte remuestreo de los clados, como se describe en la presente memoria, está indicado a continuación de sus ramas apropiadas. Además, las longitudes de las ramas están representadas a escala para las distancias genéticas.

Las cepas 23 y 25 se agrupan de nuevo como un linaje hermano para el clado HCV 1b. Estos datos junto con los datos de aminoácidos, sugieren su representación de un tercer clado genéticamente diferenciado. Las relaciones jerárquicas dentro de los dos principales clados difieren algo del árbol de consenso obtenido por Parsimony, pero ninguna de las diferencias está sólidamente basada en el remuestreo con empalme de vecinos. El clado HCV 1b está basado en el 100% de las muestras del remuestreo. El agrupamiento 23-25 separado está basado en el 97,4% de las muestras del remuestreo. Dentro del clado HCV 1a, las cepas 4 y 1Y son monofiléticas con un respaldo del 84,9%, las cepas 11, 170 177 son monofiléticas con un respaldo del 66%.

25 Análisis de las tasas evolutivas.

Las tasas poblacionales de evolución para los tres linajes se estimaron siguiendo el método de Watterson, G. A., 1975, supra. Los estimadores de θ (+ desviación típica) para HCV 1a, HCV 1b y los pacientes 23 -25 son $42,28 \pm 53,77$, $49,49 \pm 61,81$ y $40 \pm 40,50$, respectivamente. Por tanto, no hay indicación de tasas diferenciales de evolución que ocurren entre los tres linajes. Los ensayos de evolución no neutra entre los dos principales linajes HCV 1a y HCV 1b se realizaron usando el método de McDonald y Kreitman (McDonald, J. et al., (1991) Adaptive protein evolution at adh locus in *Drosophila*. Nature 351: 652-65441). Bajo neutralidad la relación de sitios de segregación sinónimos y no sinónimos debe ser constante dentro de y entre las poblaciones. En comparaciones de las poblaciones de HCV 1a y HCV 1b, el valor X^2 es 2,30 indicando que no había una desviación significativa de las expectativas neutras para la desviación de secuencias entre las dos poblaciones.

35 Modelización molecular.

Los experimentos de modelización de proteínas por homología del complejo proteasa NS3/NS4A del HCV se llevaron a cabo con el molde de la estructura cristalina de 2,2 Å de la proteasa NS3 y el péptido NS4A (Yan, Y. 1998, supra); para examinar los cambios tridimensionales debidos a variación de la secuencia en las cepas clínicas. Las cepas más divergentes se seleccionaron para modelización. El foco se puso en los subtipos A y B y, por tanto, se seleccionaron para estos estudios dos subtipos 1A altamente divergentes y dos aislados muy diferentes en cuanto a secuencia, que pertenecen al subtipo 1B. Para los experimentos de modelización se han seleccionado cuatro cepas clínicas (252, 1G, 1H y 1U). Dos de las cepas (252 y 1H) pertenecen al subtipo 1a y las otras dos cepas (1G y 1U) pertenecen al subtipo 1b. Las mutaciones ocurren más frecuentemente en las regiones de bucle o en los extremos de las cadenas β . Las mutaciones también ocurren en las α -hélices de NS3.

Los modelos de homología para las secuencias de consenso de HCV 1a o HCV 1b que usan la estructura cristalina del complejo NS3/NS4A (Yan Y. et al., 1998, supra) se trataron en ordenador con la suite de programas informáticos MODELLER. Los modelos para las secuencias de consenso 1a y 1b se superpusieron para un análisis detallado (véase la Figura 4 en el documento de Holland-Staley, C. A., 2002, supra), que indica las mayores diferencias entre los dos subtipos. Además, había diferencias apreciables tanto en el sitio activo como en las otras posiciones. Hay diferencias distinguibles en las posiciones de los tres residuos de la triada catalítica, Asp81, His57 y Ser139. Los resultados de la modelización sugieren diferencia funcionales para los dos subtipos.

Los modelos de las cepas clínicas del HCV 252, 1G, 1H y 1U que contienen múltiples cambios de aminoácidos se superpusieron sobre la estructura cristalina del complejo NS3/NS4A de tipo silvestre de Yan, Y. et al., 1998, supra. El complejo NS3/NS4A de la cepa 1G se comparó con la estructura cristalina del tipo silvestre (véase la Figura 5 del documento de Holland-Staley, C. A., 2002, supra). Hay diferencias distinguibles en las posiciones de los tres residuos para la tríada catalítica (Asp81, His57 y Ser139). Se han observado similares diferencias potenciales para las cepas 252, 1H y 1U (datos no mostrado) (Holland-Staley, C. A., et al., (2002) Genetic diversity and response to

IFN of the NS3 protease gene from clinical strains of the hepatitis C virus. Arch. Virol. 147: 1385-1406), que se incorpora en la presente memoria como referencia.

EJEMPLO 2

5 UNA DISMINUCIÓN EN LA DIVERSIDAD GENÉTICA DENTRO DEL GEN DE LA HELICASA DE NS3 PROCEDENTE DE AISLADOS CLÍNICOS DEL VIRUS DE LA HEPATITIS C ESTÁ CORRELACIONAOD CON LA TERAPIA DE INTERFERÓN.

Este ejemplo describe los efectos de la terapia con interferón (con o sin ribavirina) en el gen de la helicasa de NS3 de pacientes infectados con HCV tipo 1a o 1b. Para hacer esto, se desarrollo una PCR anidada, que permite la recuperación de datos de la secuencia de NS3 del HCV directamente de los aislados de pacientes. Para analizar los efectos de la terapia con IFN, se preparó una secuencia de consenso derivada de paciente y se usó para determinar el número global de mutaciones en cada aislado clínico. Luego las secuencias se agruparon respecto a si el paciente había tenido o no había tenido terapia con IFN (tratados vs experimentados), y se analizaron el número de posiciones que habían mutado. Estas mutaciones también se usaron para predecir la respuesta al IFN. Los resultados demuestran que ocurre una disminución significativa de la variabilidad de secuencias global en aquellos aislados que habían sido sometidos a terapia con IFN. Por tanto, la presente invención proporciona un método para analizar los efectos de IFN sobre el gen NS3 del HCV y determinar su efecto sobre la población de cuasiespecies.

MATERIALES Y MÉTODOS

Muestras de pacientes y extracción de RNA viral.

Se analizó el suero de cuarenta y tres pacientes infectados con los subtipos 1a y 1b del HCV. El estudio incluía veintitrés varones y veinte mujeres con una media de edad de 51,1 años (intervalo, 21- 76). Veinte nueve no tuvieron tratamiento, mientras que los 14 restantes habían sido tratados con interferón solo o combinado con ribavirina. Todos los pacientes tenían cargas de RNA viral del HCV de entre 5,1 x 10⁴ y 2,3 x 10⁶ viriones/ml determinadas usando el Roche Amplicor HCV Monitor Test. El RNA del HCV se aisló del suero de pacientes usando el kit de mini-aislamiento de RNA viral QIAamp (de Qiagen Inc., Valencia, CA). El aislamiento se realizó de acuerdo con las instrucciones del fabricante.

RT-PCR y amplificación de RNA viral.

La reacción en cadena de la polimerasa con transcriptasa inversa (RT-PCR) seguida por una PCR 'anidada' de 2ª ronda se usó para amplificar el gen NS3 completo del HCV subtipos 1a y 1b procedentes de aislados clínicos. La etapa de RT-PCR usó el kit Promega Access Reverse Transcriptase PCR (de Promega Corporation, Madison, WI). Para esto se diseñaron dos cebadores de oligonucleótidos que flanquean el gen. El cebador 5' (5'-CTA CTC CTG CTC CTG CTG GCG TT-3') se hibrida 693 pb antes del comienzo del gen NS3. El cebador 3' (5'-CTG ATG AAA TTC CAC ATG TGC TTC GCC CA-3') se hibrida 338 pb aguas abajo de NS3 y se usó para iniciar la etapa de la transcriptasa inversa (RT). Esto permite la producción de cDNA seguido por la amplificación inicial de la región deseada. La mezcla de la amplificación que contenía 25 pmol de cada cebador, diversos dNTP 200 µM, RT-polimerasa de AMV 2,5 U, DNA-polimerasa de Tfl 2,5U, MgSO₄ 1,5 mM y 20 µl de RNA viral se añadió a un aparato de ciclo térmico (48°C) Perkin-Elmer 9700 (de Perkin Elmer Cetus, Foster City, CA). El protocolo de la PCR consistió en una etapa de transcriptasa inversa (RT) a 48°C durante 45 minutos, seguido por una desnaturalización inicial a 94°C durante 2 minutos, y 35 ciclos de 94°C durante 15 segundos; 55°C durante 20 segundos; 72°C durante 2 minutos; y una extensión final a 72°C durante 10 minutos.

Para la segunda etapa de la PCR, el producto de la 1ª ronda de la PCR se amplificó usando cebadores, que se hibridan en el interior de la reacción previa, creando una amplificación 'anidada'. La reacción de la segunda ronda de la PCR usó el cebador 5' (5'-GAG CCC GTC GTC TTC TC-3') y el cebador 3' (5'-CAC TCT TCC ATC TCA TCG AAC TCC TGG TAG AG-3'). La mezcla de amplificación que contenía 25 pmol de cada cebador, diversos dNTP 200µM, RT-polimerasa Taq® 2,5 U, MgSO₄ 1,5 mM y 5 µl del producto de reacción de la 1ª ronda se añadió a un aparato de ciclo térmico (94°C). La desnaturalización inicial consistió en 10 minutos a 94°C, seguido por 35 ciclos de 94°C durante 15 segundos; 55°C durante 20 segundos; 72°C durante 2 minutos y una extensión final a 72°C durante 10 minutos. El producto de amplificación resultante es una sola banda de 2219 pb sobre un gel de agarosa al 1%. El gen NS3 completo gene se amplificó usando este método, sin embargo, solamente el gen de la helicasa de la NS3 se caracteriza en la presente invención. Los productos de la segunda etapa de reacción se purificaron y concentraron usando el Gene Clean Spin Protocol (de Bio 101, Vista, CA). Los productos purificados se usaron directamente para secuenciar DNA. Para asegurar que cada aislamiento de RNA del HCV era satisfactorio se usaron cebadores que cubren la región 5'-no traducida (5'-UTR) altamente conservada como control (Yuki, 1997 N°226). Los productos de PCR de esta región de control también se purificaron como antes y se secuenciaron. Los datos de secuencias de la región 5'UTR se usaron luego para determinar el genotipo de cada aislado de HCV (O'Brien, 1997 N° 30). Si la amplificación no fue satisfactoria se realizó una nueva ronda de RT-PCR usando un segundo cebador 3' "de respaldo". El cebador 3' de respaldo (5'-GCC GAC ATG CAT GYC ATG ATG TAT TT-3') da como resultado un producto de amplificación de 2039 pb, que recorta el gen NS3 en 20 pb.

Secuenciación de DNA.

La secuenciación de los productos purificados de la PCR de la segunda etapa se realizó usando la tecnología ABI Big Dye Terminator (de Applied Biosystems, Foster City, CA). Se usaron once cebadores (A-K) diseñados para cubrir la región de la helicasa NS3 tanto en las cadenas con sentido como en las cadenas antisentido. Los cebadores con sentido son : A (5'-TCG GAC CTT TAC TTG GTC ACG AG-3'), B (5'-TAC TCC ACC TAT GGC AAG TTC CT-3'), C (5'-CAT CTC ATY TTC TGC CA-3'), D (5'-GAG TGC TAT GAC GCG GGC TGT GCT T-3'), E (5'-CCA TCG TGG GAY CAA ATG TGG AAG TGT-3'), F (5'-AAA TAC ATC ATG RCA TGC ATG TCG GC-3'). Los cebadores antisentido son: G (5'-AGG AAC TTG CCA TAG GTG GAG TA-3'), H (5'-GAG TGG CAC TCA TCA CA-3'), I (5'-TCG ACT GTC TGR GTG ACA CA-3'), J (5'-GCC GAC ATG CAT GYC ATG ATG TAT TT-3'), K (5'-CAC TCT TCC ATC TCA TCG AAC TCC TGG TAG AG-3'). Para el análisis ABI, las reacciones de secuenciación se llevaron a cabo en placas de micro-valoración usando un aparato de ciclo térmico de 25 ciclos de 96°C durante 10 segundos; 50°C durante 5 segundos; y 60°C durante 4 minutos. Los productos de reacción se purificaron con NaOAc/EtOH, se volvieron a poner en suspensión en tampón de carga, se desnaturalizaron y se analizó en el secuenciador ABI 377 (de Applied Biosystems).

15 Análisis de las secuencias.

Los resultados de la secuenciación con ABI se analizaron usando el programa informático de análisis de secuencias ABI versión 3.3. Las secuencias individuales se alinearon frente a una secuencia de consenso derivada de datos de secuencias de todos los pacientes y frente a los tipos del HCV 1a (Número de acceso AF009606) y 1b (Número de acceso AJ000009) depositados en Genbank (Yanagi, 1997 N° 227) (Yanagi, 1998 N° 228). Las dos secuencias de Genbank se eligieron para comparación porque estaban disponibles como construcciones clonadas para uso como controles positivos para las valoraciones de expresión de proteínas subsiguiente (datos no mostrados). Los aminoácidos que difieren de la secuencia de consenso se comparan y usan para determinar cuáles residuos son los conservados y cuáles son los variables.

Números de acceso a GenBank de las secuencias de nucleótidos.

25 Las secuencias de nucleótidos descritas en la presente memoria han sido depositadas en la American Type Culture Collection (ATCC), 10801 University Boulevard, Manassas, VA 20110-2209, el 9 de abril de 2001 y se les ha asignado los números de acceso AF369214 a AF369263. Estos depósitos se mantendrán bajo los términos del Tratado de Budapest referente al reconocimiento internacional del depósito de microorganismos para fines del procedimiento de patentes. Estos depósitos se hicieron simplemente para comodidad de los expertos en la técnica y no se reconoce que se requiera un depósito bajo las normas del 35 U. S. C. §112.

RESULTADOS

Análisis de la secuencia del gen de la helicasa NS3 del HCV.

Se usó una PCR anidada para amplificar y secuenciar la helicasa NS3 del HCV de los sueros de 43 pacientes infectados con HCV subtipos 1a ó 1b del HCV. La región 5' no traducida conservada se usó para colocar cada aislado en su genotipo respectivo usando el método de O'Brien et al. Los resultados del genotipo mostraron que 30 secuencias eran HCV subtipo 1a y 13 eran del subtipo 1b. Estos datos son consistentes con la población de pacientes, siendo HCV subtipos 1a y 1b los más prevalentes en esta región del mundo. De los datos de la secuencia de NS3 se generó una secuencia de consenso derivada de los pacientes y se usó para determinar la variabilidad genética de los aislados de pacientes tanto para la distribución de nucleótidos como para la sustitución de aminoácidos.

Dentro de las secuencias analizadas se identificó una variación significativa en las secuencias de nucleótidos aminoácidos. Hay 1350 nucleótidos que codifican el gen de la helicasa NS3. Los datos de secuencias se analizaron para determinar la distribución de nucleótidos en cada posición cuando se compararon con la secuencia de consenso derivada de pacientes y las dos secuencias prototipos de Genbank (Genbank N° AF009606 y N° AJ000009). Para estos aislados que pertenecen al HCV subtipo 1a, 436 posiciones contenían una o más sustituciones de nucleótidos cuando se compararon con la secuencia de consenso de pacientes, dando como resultado un variación de secuencia global de 32,30%. Estas incluían 355 (26,30%) transiciones, 30 (2,22%) transversiones y 51 (3,78%) con ambas, transiciones y transversiones. El prototipo HCV 1a difería en 34 posiciones (variación de secuencia global de 2,52%), con 30 posiciones que contenían transiciones y 4 que contenían transversiones. No se detectaron inserciones o deleciones en las secuencias de pacientes o los prototipos. Para aislados que tiene el HCV subtipo 1b, 390 posiciones contenían una o más sustituciones cuando se compararon con la secuencia de consenso de pacientes, para una variación de secuencia global del 28,89%. Había 309 (22,89%) transiciones, 36 (2,67%) transversiones y 45 (3,33%) que contenían ambas, transiciones y transversiones. Había 55 posiciones en la secuencia prototipo HCV 1b que eran diferentes de la secuencia de consenso derivada de los pacientes, con una diferencia global de secuencias de 4,07%, con 47 transiciones y 8 transversiones. En ninguna de las secuencias se detectaron inserciones o deleciones. Puesto que la tasa de mutación predicha introducida durante

la amplificación por PCR es ~0,1%, se cree que los datos presentados en la presente memoria muestran una representación precisa de la variación de la secuencia genética natural.

Hay 450 aminoácidos que comprenden la proteasa NS3. Los efectos de la sustitución de nucleótidos en la secuencia de aminoácidos mostraron que la mayoría de las sustituciones eran silenciosas. Para las secuencias de 30 pacientes que pertenecen al HCV subtipo 1a, 65 residuos o 14,44% diferían de la secuencia de consenso derivada de los pacientes (Tabla 5A). La secuencia publicada difería de la secuencia de consenso de los pacientes en 5 residuos. La mayoría de los sitios tenía una sola sustitución de especie, 6 sitios tenían tres especies de aminoácidos presentes, con 2 sitios (residuos 334 y 383; numerando dentro de NS3) capaces de tolerar 4 residuos diferentes. Incluyendo la secuencia de control, 38 sitios tenían sustituciones del mismo tipo de aminoácidos, 20 sitios eran capaces de acomodar aminoácidos polares o neutros, 2 sitios básicos o polares, 2 sitios ácidos o polares, 1 sitio básico o neutro, 1 sitio ácido o neutro, y 1 sitio básico, polar o neutro. Solamente 1 paciente (paciente 186) tenía una secuencia idéntica a la de consenso sin sustituciones de aminoácidos. La Tabla 5A resume las mutaciones que ocurren en más de una secuencia de pacientes.

Para aislados de pacientes de los tipos HCV 1b, hubo una variación de secuencia global de 11,33% con 51 sitios diferentes de la secuencia de consenso derivada de los pacientes (Tabla 5B). La secuencia de consenso de los pacientes difería de la secuencia de GenBank en 13 residuos, incluyendo 4 nuevos residuos, que llevan el número total de sitios que contienen uno o más aminoácidos a 55 o 12,22% de variación de secuencia. Tres especies de aminoácidos estaban presentes en 9 sitios, con el residuo 402 capaz de tener 5 aminoácidos diferentes y el residuo 470 capaz de tener 7 aminoácidos diferentes. El mismo tipo de sustitución de aminoácidos se vio en 26 sitios, mientras que 20 sitios eran capaces de acomodar aminoácidos neutros o polares, 5 sitios básicos o polares, 1 sitio ácido o polar, 1 sitio ácido o básico y dos sitios podían acomodar aminoácidos neutros polares o básicos (Tabla 5B). Ninguna de las secuencias de pacientes fue idéntica a la secuencia de consenso.

La diferencia entre las secuencias de consenso para los tipos HCV 1a y 1b mostró una similitud de secuencia de aminoácidos de 96,4%. Los restos conservados fueron sorprendentemente similares difiriendo en solo 4 residuos (259, 263, 269 y 270). No pudo establecerse una correlación entre las mutaciones en los HCV tipos 1a y 1b. Entre todas las secuencias, no se encontraron sustituciones de aminoácidos en los residuos que son homólogos en función a las de otras helicasas de la misma familia (residuos sombreados en las Tablas 5A y 5B). Sin embargo, se encontraron varias mutaciones muy próximas a los siete restos. La mayoría de las sustituciones ocurren solamente en una secuencia, sugiriendo una sustitución aleatoria debida a la falta de incorporación de una base por la RT-polimerasa. Las secuencias del HCV 1a no mostraron sustituciones en el resto I, la sustitución de L226I en el resto 1a (paciente 4C), 1 sustitución de S263G en el ácido nucleico que une el resto (paciente 4B), 1 sustitución V329I en el resto III (pacientes. 23, 24, 270, 1H, IX, 3T, 1C), 2 sustituciones en el resto IVa; K371R (paciente ID) y K372R (pacientes 4 y 1C), y 3 sustituciones en el resto V; A410S (control 1a), A410T (paciente 3S) y F418Y (pacientes 4, 12, 177, 194, 198, 3P, 3S, 3T, 4C, K y 174). El primer resto guardián tenía 1 sustitución, S424T (pacientes 1X y K), el resto VI tenía 5 sustituciones; R458S (paciente 3S), T459S (paciente 5P), R461L (paciente 3S), K469R (pacientes 1X, 5P, 174 y 183) y P470Q (paciente 4C), y el último resto guardián tenía 3 sustituciones; D496N (paciente 38), A497T (paciente 11) y G498S (paciente 4B). De todas las mutaciones, las sustituciones en el resto IVa fueron las más sorprendentes puestos que estos dos residuos de lisina, en las posiciones 371 y 372, han demostrado ser importantes es estabilizar la cadena de ácido nucleico monocatenario. El cambio de aminoácidos de lisina a arginina en ambos casos indica la necesidad de un aminoácido básico en esta posición. Otra sustitución que fue interesante estaba en la posición 410 en el resto V. Esto fue interesante porque la secuencia prototipo contenía una serina en lugar de alanina, una serina en esta posición está en la mayoría de las secuencias publicadas, sin embargo, la secuencia presentada en la invención identificada como SEC ID NO: 1 mostró una presencia de alanina en la secuencia de consenso derivada de pacientes excepto para un paciente que tenía una treonina en esta posición (paciente 3S). Las posiciones 229, 418 y 469 demostraron ser interesantes porque varios pacientes contenían estas sustituciones, sin embargo, el significado de las mismas sigue desconocido.

Las sustituciones de aminoácidos en aislados de HCV 1a en las posiciones de los restos existentes exteriores mostró una alta prevalencia de mutaciones en varias regiones. Un cambio de isoleucina a valina en la posición 248 ocurrió en 9 de 30 pacientes con HCV 1a. Los residuos circundantes desde 240 a 252 también mostraron un alto grado de mutación. El residuo 281 podía ser glicina (consenso) o arginina (7 pacientes). Esto es sorprendente porque la transición es desde una cadena lateral sencilla a una cadena lateral voluminosa. Otras regiones que contienen mayor frecuencia de sustitución fueron desde el residuo 229 al 358 al 382 al 386, los residuos 455 a 461, el residuo 557 y el extremo COOH.

El resto I en secuencias de HCV 1b tenía 1 sustitución, K213R (pacientes 3N y 3Q), ninguna sustitución en 1a, 3 en el resto de unión de ácido nucleico; A263G (control), P264S (control, pacientes 1A, 2D, 3N, 6A) y I265V (paciente 3Q). El resto II tenía 2 sustituciones; I288M (paciente 1U), y T295I (paciente 1A), los restos III and IVa no contenían mutaciones, mientras que el resto V contenía 2; V406A (paciente. 30) y F418Y (paciente 3Q). Ambos restos guardianes no contenían mutaciones. El resto VI contenía 1 en la posición 470. Esta mutación era significativa porque se encontraron un total de 7 aminoácidos: R470S (control y paciente 6A), R470M (pacientes 1G, 1U y 3N), R470G (pacientes 179 y 3Q), R470V (paciente 28), R470P (paciente 30) y R470A (paciente 1). Similitudes al HCV

1a se encontraron en regiones fuera de los restos conocidos. Las posiciones 240 a 256, 334 a 358, 382 a 386 y el extremo o término COOH contenían regiones “calientes”.

Los efectos del tratamiento con IFN.

5 Las diferencias más significativas se apreciaron cuando los aislados se agruparon según si los pacientes habían recibido o no habían recibido terapia con interferón. Se puso de manifiesto una disminución global significativa en la variabilidad de las secuencias en ambos tipos de HCV 1a y 1b después de la terapia con IFN con o sin ribavirina. Por tanto, el interferón desempeña un papel en la selección de cuasiespecies.

10 Para aislados de HCV 1a que no habían sido sometidos a terapia con interferón el número medio global de mutaciones por paciente fue 6,47 por secuencia. Un total de 59 residuos diferentes estaban afectados. Después de la terapia con interferón, el número medio de mutaciones fue 5,0 por paciente, conteniendo 25 sitios aminoácidos diferentes lo que da una caída global de 56,14% en posiciones mutadas (Tabla 5A). Para los aislados del tipo 1b, el número medio de mutaciones por paciente para los aislados no sometidos a terapia con interferón fue 8,9, con una variación global de 53 sitios. Sin embargo, después de la terapia con interferón, el número de mutaciones por paciente fue 6,33, y el número de sitios que contenían mutaciones descendió espectacularmente en los aislados 1b
15 a solamente 16, o en 69,78% (Tabla 5B). Esto sugiere que las cepas sensibles fueron eliminadas durante la terapia con interferón y estos sitios restantes pueden estar implicados de algún modo en la resistencia de los virus al interferón. No se apreciaron diferencias en el tipo 1a o 1b para estos aislados que fueron sometidos a una terapia de combinación con ribavirina frente a los que fueron tratados con solo interferón, y no pudo establecerse una correlación entre el número o posiciones de las mutaciones y la carga viral.

20

LISTA DE SECUENCIAS

5	<110> Holland-Staley, Carol	
	<120> MÉTODOS Y COMPOSICIONES PARA IDENTIFICAR Y CARACTERIZAR LA HEPATITIS C	
	<130> PCV-001PC	
10	<150> 60/363,603	
	<151> 2002-03-11	
	<160> 64	
15	<170> FastSEC for Windows Versión 4.0	
	<210> 1	
20	<211> 17	
	<212> DNA	
	<213> Secuencia Artificial	
	<220>	
25	<223> Cebador	
	<400> 1	
	cctgcttgat gatgatg	17
30	<210> 2	
	<211> 29	
	<212> DNA	
	<213> Secuencia Artificial	
35	<220>	
	<223> Cebador	
	<400> 2	
40	ctgatgaaat tccacatggt cttcgccca	29
	<210> 3	
	<211> 23	
	<212> DNA	
	<213> Secuencia Artificial	
45	<220>	
	<223> Cebador	
	<400> 3	
50	ctactcctgc tctgctggc gtt	23
	<210> 4	
	<211> 32	
	<212> DNA	
55	<213> Secuencia Artificial	
	<220>	
	<223> Cebador	
60	<400> 4	
	cactcttcca tctcatcgaa ctctggtag ag	32
	<210> 5	
	<211> 21	

<212> DNA
 <213> Secuencia Artificial

 <220>
 5 <223> Cebador

 <400> 5
 gtgtgggtyc cccccctcaa c 21

 10 <210> 6
 <211> 26
 <212> DNA
 <213> Secuencia Artificial

 15 <220>
 <223> Cebador

 <400> 6
 gccgacatgc atgycatgat gtatt 26
 20
 <210> 7
 <211> 17
 <212> DNA
 <213> Secuencia Artificial
 25
 <220>
 <223> Cebador

 <400> 7
 30 gagcccgtcg tcttctc 17

 <210> 8
 <211> 17
 <212> DNA
 35 <213> Secuencia Artificial

 <220>
 <223> Cebador

 <400> 8
 40 ggccgggaca asaacca 17

 <210> 9
 <211> 23
 <212> DNA
 45 <213> Secuencia Artificial

 <220>
 <223> Cebador
 50
 <400> 9
 tcggacctt acttggcac gag 23

 <210> 10
 <211> 23
 <212> DNA
 <213> Secuencia Artificial

 <220>
 60 <223> Cebador

 <400> 10
 aggaactgc cwtagggtga gta 23

<210> 11
 <211> 23
 <212> DNA
 <213> Secuencia Artificial
 5
 <220>
 <223> Cebador
 <400> 11
 10 cgccggcgca ccggaatgac atc 23
 <210> 12
 <211> 23
 <212> DNA
 15 <213> Secuencia Artificial
 <220>
 <223> Cebador
 20 <400> 12
 tactccacct atggcaagtt cct 23
 <210> 13
 <211> 25
 25 <212> DNA
 <213> Secuencia Artificial
 <220>
 <223> Cebador
 30 <400> 13
 gactgctatg acgcgggctg tgctt 25
 <210> 14
 <211> 17
 <212> DNA
 <213> Secuencia Artificial
 35 <220>
 <223> Cebador
 40 <400> 14
 gactggcact catcaca 17
 45 <210> 15
 <211> 17
 <212> DNA
 <213> Secuencia Artificial
 50 <220>
 <223> Cebador
 <400> 15
 55 catctcatyt tctgccca 17
 <210> 16
 <211> 20
 <212> DNA
 <213> Secuencia Artificial
 60 <220>
 <223> Cebador
 <400> 16
 65 tcgactgtct grgtgacaca 20

<210> 17
 <211> 27
 <212> DNA
 5 <213> Secuencia Artificial

 <220>
 <223> Cebador

 10 <400> 17
 ccatcgtggg aycaaatgtg gaagtgt 27

 <210> 18
 <211> 26
 15 <212> DNA
 <213> Secuencia Artificial

 <220>
 <223> Cebador
 20
 <400> 18
 aaatacatca tgrcatgcat gtcggc 26

 <210> 19
 <211> 29
 25 <212> DNA
 <213> Secuencia Artificial

 <220>
 <223> Cebador
 30
 <400> 19
 tgggcgaagc acatgtggaa tttcatcag 29

 <210> 20
 <211> 23
 35 <212> DNA
 <213> Secuencia Artificial

 <220>
 <223> Cebador
 40
 <400> 20
 gccatcctct ctctggtgc cct 23
 45
 <210> 21
 <211> 21
 <212> DNA
 <213> Secuencia Artificial
 50
 <220>
 <223> Cebador

 <400> 21
 55 cacccayccc cccaakatgt t 21

 <210> 22
 <211> 24
 <212> DNA
 60 <213> Secuencia Artificial

 <220>
 <223> Cebador

	<400> 22 ccggrggtca ttaygtscaa atgg	24
5	<210> 23 <211> 18 <212> DNA <213> Secuencia Artificial	
10	<220> <223> Cebador	
	<400> 23 cgcggggttt ccaggcag	18
15	<210> 24 <211> 20 <212> DNA <213> Secuencia Artificial	
20	<220> <223> Cebador	
	<400> 24 ggcacytaca tctatgayca	20
25	<210> 25 <211> 21 <212> DNA <213> Secuencia Artificial	
30	<220> <223> Cebador	
	<400> 25 tcatcttcag tccgatggag a	21
35	<210> 26 <211> 23 <212> DNA <213> Secuencia Artificial	
40	<220> <223> Cebador	
	<400> 26 atgccgctra tgaartcca cat	23
45	<210> 27 <211> 22 <212> DNA <213> Secuencia Artificial	
50	<220> <223> Cebador	
	<400> 27 ctgaaygccca tcatggargc ca	22
55	<210> 28 <211> 19 <212> DNA <213> Secuencia Artificial	
60	<220> <223> Cebador	
65	<210> 28 <211> 19 <212> DNA <213> Secuencia Artificial	
	<220> <223> Cebador	

ES 2 386 742 T3

	<400> 28 ccaatggaga agaaggtca	19
5	<210> 29 <211> 18 <212> DNA <213> Secuencia Artificial	
10	<220> <223> Cebador	
15	<400> 29 atgtggagac atcctgca	18
20	<210> 30 <211> 18 <212> DNA <213> Secuencia Artificial	
	<220> <223> Cebador	
25	<400> 30 ttgtggcctg ttgtagga	18
30	<210> 31 <211> 18 <212> DNA <213> Secuencia Artificial	
	<220> <223> Cebador	
35	<400> 31 gcgcattctt ccatctca	18
40	<210> 32 <211> 18 <212> DNA <213> Secuencia Artificial	
45	<220> <223> Cebador	
50	<400> 32 ccgaaacgct gaygtcat	18
55	<210> 33 <211> 18 <212> DNA <213> Secuencia Artificial	
	<220> <223> Cebador	
60	<400> 33 ggatggaggc tgctcagc	18
	<210> 34 <211> 18 <212> DNA <213> Secuencia Artificial	

	<220> <223> Cebador	
5	<400> 34 agacccgacc ttaccat	18
10	<210> 35 <211> 20 <212> DNA <213> Secuencia Artificial	
15	<220> <223> Cebador <400> 35 ttaccctgt gtcaagacca	20
20	<210> 36 <211> 18 <212> DNA <213> Secuencia Artificial	
25	<220> <223> Cebador <400> 36 ccaatgatgg aratgcag	18
30	<210> 37 <211> 18 <212> DNA <213> Secuencia Artificial	
35	<220> <223> Cebador <400> 37 ccttgcct agcgcata	18
40	<210> 38 <211> 18 <212> DNA <213> Secuencia Artificial	
45	<220> <223> Cebador <400> 38 gagcactacg ctgtcgaa	18
50	<210> 39 <211> 18 <212> DNA <213> Secuencia Artificial	
55	<220> <223> Cebador <400> 39 tggtgtggc tagaacca	18
60	<210> 40 <211> 18 <212> DNA <213> Secuencia Artificial	
65	<210> 40 <211> 18 <212> DNA <213> Secuencia Artificial	

<220>
 <223> Cebador
 5 <400> 40
 gtaggcccga aaaccaag 18
 <210> 41
 <211> 18
 10 <212> DNA
 <213> Secuencia Artificial
 <220>
 <223> Cebador
 15 <400> 41
 gcyctgaaca agcccacg 18
 <210> 42
 <211> 18
 20 <212> DNA
 <213> Secuencia Artificial
 <220>
 <223> Cebador
 25 <400> 42
 acatctgrgt gactggtc 18
 <210> 43
 <211> 19
 30 <212> DNA
 <213> Secuencia Artificial
 35 <220>
 <223> Cebador
 <400> 43
 ccygtaatat ttagtccca 19
 40 <210> 44
 <211> 18
 <212> DNA
 <213> Secuencia Artificial
 45 <220>
 <223> Cebador
 <400> 44
 50 tggaaatcaa ggtcatca 18
 <210> 45
 <211> 19
 <212> DNA
 55 <213> Secuencia Artificial
 <220>
 <223> Cebador
 60 <400> 45
 gtagctacta tgggctcaa 19
 <210> 46

<211> 18
 <212> DNA
 <213> Secuencia Artificial

5 <220>
 <223> Cebador

<400> 46
 gcttgctcga tgtacggg 18

10 <210> 47
 <211> 19
 <212> DNA
 <213> Secuencia Artificial

15 <220>
 <223> Cebador

20 <400> 47
 cccgtcatct ttagtccta 19

25 <210> 48
 <211> 18
 <212> DNA
 <213> Secuencia Artificial

<220>
 <223> Cebador

30 <400> 48
 tggagattaa gggtatca 18

35 <210> 49
 <211> 19
 <212> DNA
 <213> Secuencia Artificial

<220>
 <223> Cebador

40 <400> 49
 gattgcacta tgggtcgaa 19

45 <210> 50
 <211> 18
 <212> DNA
 <213> Secuencia Artificial

50 <220>
 <223> Cebador

<400> 50
 gcttgctcga tgtaagga 18

55 <210> 51
 <211> 17
 <212> DNA
 <213> Secuencia Artificial

60 <220>
 <223> Cebador

<400> 51
 ggccgggaca asaacca 17

65

<210> 52
 <211> 23
 <212> DNA
 <213> Secuencia Artificial
 5
 <220>
 <223> Cebador
 <400> 52
 10 tcggaccttt acttggtcac gag 23
 <210> 53
 <211> 23
 <212> DNA
 15 <213> Secuencia Artificial
 <220>
 <223> Cebador
 20 <400> 53
 cgccggcgcga ccggaatgac atc 23
 <210> 54
 <211> 23
 25 <212> DNA
 <213> Secuencia Artificial
 <220>
 <223> Cebador
 30 <400> 54
 tactccacct atggcaagtt cct 23
 <210> 55
 <211> 23
 35 <212> DNA
 <213> Secuencia Artificial
 <220>
 40 <223> Cebador
 <400> 55
 aggaactgc cwtaggtgga gta 23
 45 <210> 56
 <211> 17
 <212> DNA
 <213> Secuencia Artificial
 50 <220>
 <223> Cebador
 <400> 56
 55 gagtggcact catcaca 17
 <210> 57
 <211> 17
 <212> DNA
 <213> Secuencia Artificial
 60 <220>
 <223> Cebador
 <400> 57
 65 catctcatyt tctgcca 17

<210> 58
 <211> 20
 <212> DNA
 5 <213> Secuencia Artificial

 <220>
 <223> Cebador

 10 <400> 58
 tcgactgtct grtgacaca 20

 <210> 59
 <211> 25
 15 <212> DNA
 <213> Secuencia Artificial

 <220>
 <223> Cebador
 20
 <400> 59
 gaggctatg acgcgggctg tgctt 25

 <210> 60
 <211> 27
 <212> DNA
 <213> Secuencia Artificial

 <220>
 <223> Cebador
 30
 <400> 60
 ccatcgtggg aycaaatgtg gaagtgt 27

 35 <210> 61
 <211> 26
 <212> DNA
 <213> Secuencia Artificial

 <220>
 <223> Cebador
 40
 <400> 61
 aaatacatca tgrcatgcat gtcggc 26
 45
 <210> 62
 <211> 26
 <212> DNA
 <213> Secuencia Artificial
 50
 <220>
 <223> Cebador

 <400> 62
 55 gccgacatgc atgycatgat gtattt 26

 <210> 63
 <211> 32
 <212> DNA
 60 <213> Secuencia Artificial

 <220>
 <223> Cebador

ES 2 386 742 T3

<400> 63
cactcttcca tctcatcgaa ctctggtag ag

32

5 <210> 64
<211> 23
<212> DNA
<213> Secuencia Artificial

10 <220>
<223> Cebador

<400> 64
gtgtgggttc cccccctcaa cgt

23

15

20

REIVINDICACIONES

1. Una molécula de ácido nucleico que tiene una secuencia de nucleótidos que comprende una secuencia seleccionada del grupo que consiste en:
- 5
- a) un ácido nucleico que tiene una secuencia de nucleótidos de SEC ID NO: 3;
 - b) un ácido nucleico que tiene una secuencia de nucleótidos de la SEC ID NO: 3 en donde el ácido nucleico es capaz de asociarse al gen NS4 del HCV o una de sus porciones;
 - c) un fragmento de al menos 8 nucleótidos de un ácido nucleico que tiene una secuencia de nucleótidos de SEC ID NO: 3; y
 - d) un ácido nucleico que es al menos 80% idéntico a una secuencia de nucleótidos como se ha definido en uno cualquiera de (a) a (c),
- 10
2. La molécula de ácido nucleico de la reivindicación 1, que además comprende un marcador seleccionado del grupo que consiste en un grupo fluorescente, digoxigenina, biotina, marcadores radiactivos, grupos quimioluminiscentes, enzimas, anticuerpos, agentes luminiscentes, agentes precipitantes y colorantes.
- 15
3. Un oligonucleótido para amplificar una secuencia ácido nucleico en una muestra, en donde la muestra proviene de un paciente infectado con el HCV y que contiene una secuencia de nucleótidos que consiste en la SEC ID NO: 3 o uno de sus complementos.
- 20
4. Un kit para la detección del HCV, que comprende un oligonucleótido o un conjunto de oligonucleótidos, en donde el oligonucleótido comprende la molécula de ácido nucleico de la reivindicación 1.
5. El kit de la reivindicación 4, en donde dicho kit comprende además uno o más componentes seleccionados del grupo que consiste en: al menos un oligonucleótido marcado para detectar ácido nucleico del HCV amplificado; una pluralidad de nucleótidos, una polimerasa de ácido nucleico, al menos otro componente para realizar una reacción de amplificación por polimerasa e instrucciones para su uso.
- 25
6. Un método para amplificar un ácido nucleico diana, comprendiendo el método:
- combinar un ácido nucleico diana proveniente del genoma del HCV en condiciones que permitan que ocurra una reacción de amplificación con:
- a) una o más secuencias de cebador de ácido nucleico que son al menos 80% idénticas a la secuencia descrita como SEC ID NO: 3;
 - b) una polimerasa de ácido nucleico; y
 - c) una pluralidad de nucleótidos, que dan como resultado un ácido nucleico diana amplificado.
- 30
7. El método de la reivindicación 6, en el que el ácido nucleico diana es del gen NS3 o NS4 del genoma del HCV.
8. Un método para detectar específicamente ácidos nucleicos de HCV en una muestra que comprende:
- 35
- a) poner en contacto dicha muestra con una o más secuencias de ácido nucleico que son al menos 80% idénticas a las secuencias descritas como SEC ID NO: 3, en condiciones tales que dichos ácidos nucleicos del HCV pueden hibridarse con dichos cebadores;
 - b) someter a transcripción inversa y amplificación a dichos ácidos nucleicos para obtener ácidos nucleicos del HCV amplificados; y
 - c) detectar la presencia de dichos ácidos nucleicos del HCV amplificados.
- 40
9. El método de la reivindicación 8, en donde la detección de la presencia de dichos ácidos nucleicos del HCV amplificados comprende:
- a) poner en contacto dichos ácidos nucleicos del HCV amplificados con un oligonucleótido marcado para obtener ácidos nucleicos del HCV marcados; y
 - b) identificar dichos ácidos nucleicos marcados.
- 45

10. El método de una cualquiera de las reivindicaciones 6 a 9, que comprende además la etapa de secuenciar el ácido nucleico amplificado, y opcionalmente la etapa de evaluar la secuencia respecto a mutaciones.

11. Un método in vitro para determinar si un paciente es resistente a un agente, comprendiendo el método:

5 a) realizar la amplificación por PCR del DNA de una muestra que contiene DNA de un paciente usando una secuencia de cebador de ácido nucleico que es al menos 80% idéntica a la secuencia de nucleótidos de la SEC ID NO: 3; y

b) analizar el producto amplificado, determinando con ello si un paciente es resistente a un agente.

10 12. El método de la reivindicación 11, en donde el agente es un agente antiviral, un inhibidor de proteasa, un inhibidor de la porción de serina-proteasa del gen NS3, un inhibidor de la porción de helicasa del gen NS3, un alfa-interferón, un interferón pegilado o una combinación de agentes.

13. Un método in vitro para determinar si un agente puede ser usado o no puede ser usado para tratar un paciente que tiene infección por HCV, comprendiendo el método las etapas de:

15 a) amplificar una muestra que contiene un ácido nucleico de un paciente infectado con HCV usando una secuencia de cebador de ácido nucleico que es al menos 80% idéntica a la secuencia de nucleótidos de la SEC ID NO: 3 y

b) secuenciar las secuencias de ácidos nucleicos del HCV amplificado resultante; y

c) identificar las mutaciones en la secuencia de ácido nucleico del HCV amplificado que correlacionan la resistencia o la sensibilidad a un agente antiviral, determinando de este modo si un agente puede usarse o no para tratar un paciente que tiene infección por HCV.

20 14. Un método in vitro para determinar si un tratamiento con un agente debe ser continuado o no debe ser continuado en un paciente infectado con HCV, comprendiendo el método:

a) realizar el método de la reivindicación 6 sobre dos o más muestras que comprenden DNA de un paciente durante el curso del tratamiento;

b) secuenciar el producto de ácido nucleico diana amplificado resultante de la reivindicación 6;

25 c) identificar las mutaciones presentes en el producto amplificado que están correlacionadas con la resistencia o la sensibilidad al agente; y

d) continuar el tratamiento cuando las mutaciones identificadas en el producto amplificado no cambian durante el curso del tratamiento o discontinuar el tratamiento cuando las mutaciones identificadas en el producto amplificado cambian durante el curso del tratamiento.

30 15. Un método in vitro para determinar la eficacia de un agente para tratar el HCV, comprendiendo el método:

a) secuencias de ácido nucleico de una primera muestra de un paciente expuesto a un agente, en donde las secuencias de ácidos nucleico son productos amplificados de una o más secuencias de cebador de ácido nucleico que son al menos 80% idénticas a la secuencia que se describe como la SEC ID NO: 3; y

35 b) secuencias de ácido nucleico de una segunda muestra de un paciente que no ha sido expuesto a un agente, en donde las secuencias de ácidos nucleicos son productos amplificados de una o más secuencias de cebadores de ácidos nucleicos que son al menos 80% idénticas a la secuencia que se describe como la SEC ID NO: 3;

en donde un número creciente de mutaciones en las secuencias de ácidos nucleicos de la primera muestra, con respecto a la segunda muestra, es una indicación de que el agente no es eficaz para tratar el HCV.

40

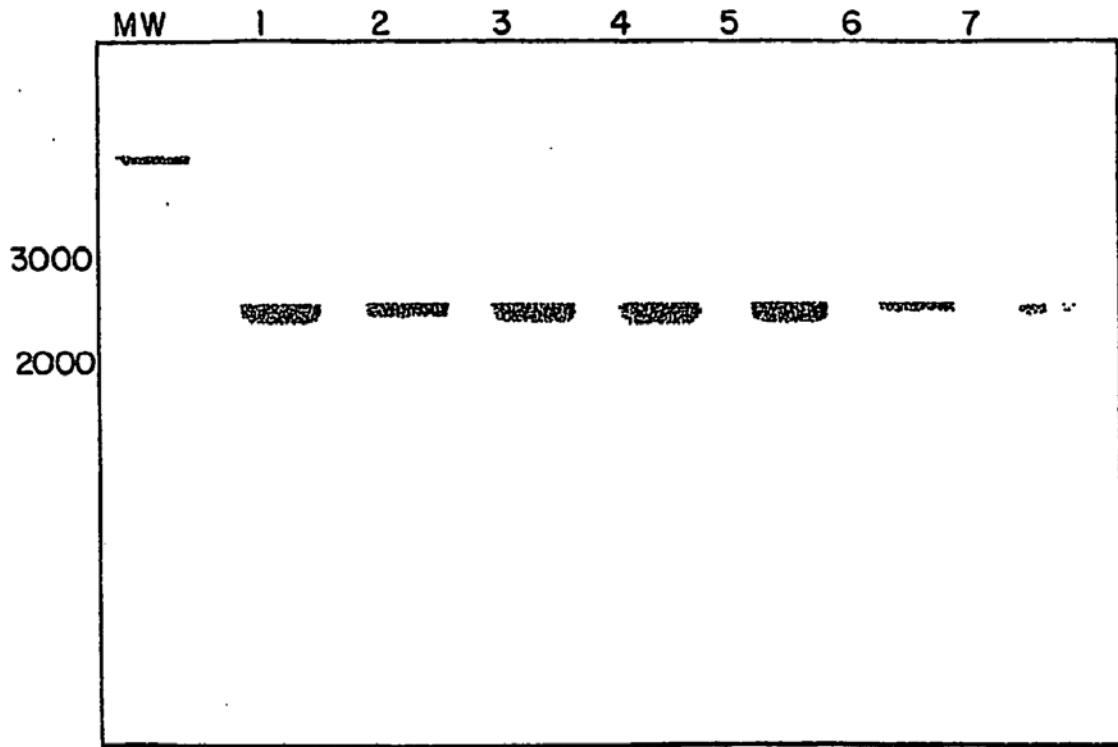


Fig. 1

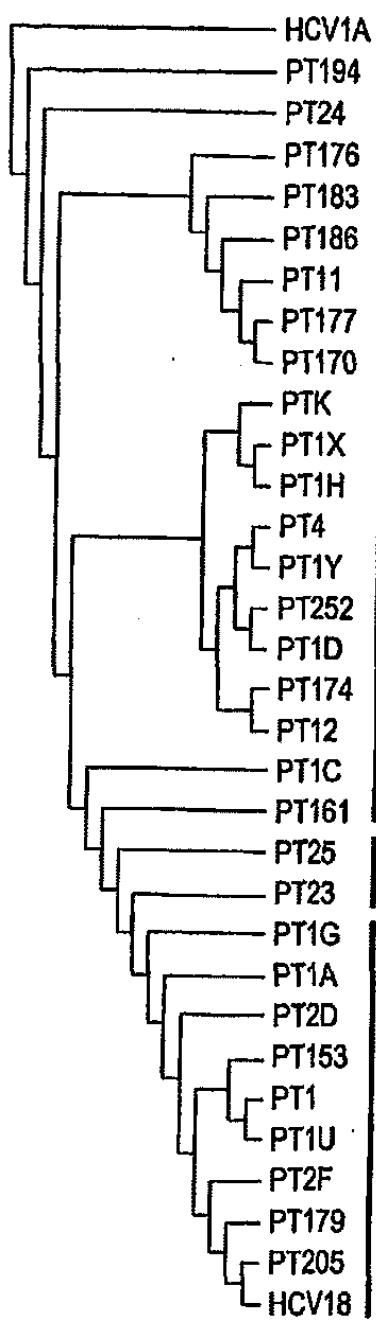


Fig. 2A

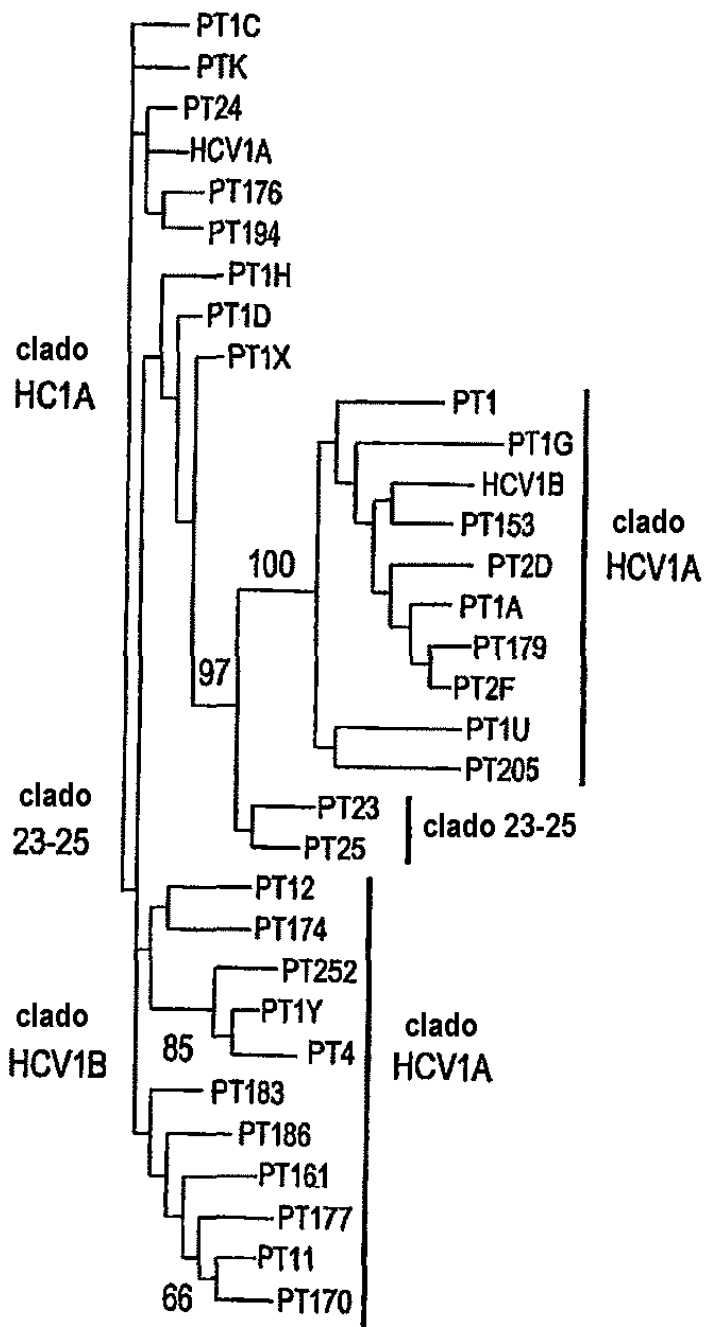


Fig. 2B