

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 676 245**

51 Int. Cl.:

C12Q 1/70

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **04.07.2014 PCT/EP2014/064281**

87 Fecha y número de publicación internacional: **08.01.2015 WO15001068**

96 Fecha de presentación y número de la solicitud europea: **04.07.2014 E 14736385 (7)**

97 Fecha y número de publicación de la concesión europea: **18.04.2018 EP 3017068**

54 Título: **Cebadores y procedimientos para detectar variantes del virus de la hepatitis C (VHC) humano en una muestra aislada**

30 Prioridad:

05.07.2013 EP 13382278

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

18.07.2018

73 Titular/es:

**FUNDACIÓ HOSPITAL UNIVERSITARI VALL D' HEBRON - INSTITUT DE RECERCA (33.3%)
Passeig Vall d'Hebrón, 119-129
08035 Barcelona, ES;
CIBER CENTRO DE INVESTIGACIÓN BIOMÉDICA EN RED (33.3%) y
F. HOFFMANN-LA ROCHE AG (33.3%)**

72 Inventor/es:

**QUER, JOSEP;
RODRÍGUEZ, FRANCISCO y
GREGORI, JOSEP**

74 Agente/Representante:

LINAGE GONZÁLEZ, Rafael

ES 2 676 245 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

Cebadores y procedimientos para detectar variantes del virus de la hepatitis C (VHC) humano en una muestra aislada

5

La presente invención se refiere al campo de la medicina en general y, específicamente, a procedimientos de genotipado para diagnóstico. En particular, la invención proporciona cebadores y procedimientos para detectar y analizar variantes de secuencia del virus de la hepatitis C (VHC) humano.

10 TÉCNICA ANTERIOR

La hepatitis C es una infección vírica del hígado que también se ha denominado "hepatitis no A, no B" (HNANB) hasta la identificación del agente causal. El virus de la hepatitis C es uno de los virus (A, B, C, D y E), que juntos representan la mayoría de los casos de hepatitis vírica. El virus de la hepatitis C se identificó por primera vez en 15 1989 y se definió como una causa frecuente de enfermedad hepática con aproximadamente 170 millones de personas infectadas en todo el mundo. La infección por el virus de la hepatitis C (VHC) afecta al hígado, lo que causa hepatitis, es decir, una inflamación del hígado. Entre un 75 y un 85 % de las personas infectadas por el VHC evolucionan a infección crónica, aproximadamente un 20 % de estos casos desarrollan complicaciones de la hepatitis C crónica, incluida cirrosis hepática o carcinoma hepatocelular después de 20 años de infección. El 20 tratamiento actual recomendado contra las infecciones por el VHC incluye una combinación de fármacos con interferón y ribavirina, añadiéndose boceprevir o telaprevir en algunos casos. En general, un 50-80 % de las personas tratadas se curan. Aquellos que desarrollan cirrosis o cáncer de hígado pueden requerir un trasplante de hígado. La hepatitis C es la principal causa de trasplante de hígado, aunque el virus generalmente reaparece después del trasplante. No hay vacuna contra la hepatitis C disponible.

25

El VHC es un virus de ARN con envoltura monocatenario de sentido (+) en el género *Hepacivirus* dentro de la familia *Flaviviridae*. El genoma vírico tiene aproximadamente 10 kilobases (kb) de longitud y codifica un precursor de poliproteína de 3011 aminoácidos. El genoma del VHC tiene un gran marco de lectura abierto (ORF) único que 30 codifica una poliproteína única, siendo dicha poliproteína procesada de forma co- y postraduccional por proteasas celulares y víricas para dar lugar a tres proteínas estructurales, es decir, núcleo, E1 (envoltura) y E2 (envoltura) y al menos siete proteínas no estructurales; p7 (canal de iones), NS2 (proteasa), NS3 (serina-proteasa/helicasa), NS4A (cofactor), NS4B (factor del complejo de replicación), NS5A (proteína resistente al interferón) y NS5B (ARN-polimerasa dependiente de ARN).

35

El VHC muestra una alta variabilidad genética. La razón de esta gran variación es una alta tasa de mutación y un alto nivel de replicación vírica mediante una ARN-polimerasa propensa a errores sin capacidad de corrección. El análisis de conjuntos extensos de secuencias de cepas aisladas del VHC en todo el mundo ha revelado la existencia de seis grupos genéticos o genotipos principales y un gran número de subtipos (también denominados subgenotipos) dentro de los seis genotipos principales. Los genotipos se numeran del 1 al 6 y los subtipos se designan como a, b, c, etc. (es decir: 1a, 1b, 2a, 2b, etc.) en ambos casos por orden de descubrimiento. La 40 divergencia general en la secuencia entre genotipos varía de un 31 a un 34 %, y de un 20 a un 23 % entre subtipos. En el documento de Simmonds et al., "Consensus proposals for a unified system of nomenclature of hepatitis C virus genotypes", *Hepatology* - 2005, vol. 42, pp. 963-973, se puede ver un extenso análisis de la base adoptada para la clasificación de genotipos y subtipos, así como los procedimientos de referencia estándar para el genotipado.

45

Determinar la infección correcta en términos del genotipo infeccioso (o incluso el subtipo) es fundamental para asignar un tratamiento correcto. De hecho, los pacientes infectados por los genotipos 1 y 4 no responden tan eficazmente al tratamiento de referencia como los otros genotipos, lo que implica un tratamiento más prolongado (48 semanas frente a 24 semanas). En cambio, por ejemplo, los genotipos 2 y 3 tienen grandes probabilidades de respuesta al tratamiento con interferón alfa en comparación con el genotipo 1. Además, cada genotipo tiene una progresión o desarrollo diferente en el huésped. De esta forma, el genotipo 1b tiene un desarrollo acelerado a cirrosis en comparación con otros genotipos. El genotipo de un sujeto infectado no cambia durante la evolución de la infección a menos que se produzca una infección adicional por otro genotipo o subtipo.

55

Hay varios kits comerciales que se utilizan para genotipar a los pacientes infectados por el VHC, incluso kits que identifican el subtipo de un genotipo. Se diferencian en el fragmento del genoma del virus que se analiza, así como en las herramientas de biología molecular empleadas en la determinación de la variante del VHC.

60

Uno de los procedimientos para el genotipado del VHC es el ensayo de hibridación inversa para el VHC Inno-Lipa (Innogenetics). En este ensayo, la región 5'-UTR del virus se analiza por hibridación inversa. Por lo tanto, los segmentos del extremo 5' no codificantes se amplifican e hibridan adicionalmente con sondas de referencia que permiten la determinación del genotipo. Una versión más robusta de Inno-Lipa es el LIPA 2.0, en el que se amplifica e hibrida una región adicional del virus, a saber, la región central codificante.

65

Otra opción es el ensayo de genotipado TruGene HCV 5'NC (Bayer). Mediante esta técnica, un segmento no codificante de la región 5' se secuencia de manera semiautomática. Las regiones 5'-UTR se amplifican, secuencian y

alinean con un panel de secuencias de referencia.

También el procedimiento de Abbott conocido como ensayo ASR para el genotipado del VHC puede ser una opción. En este caso, se realiza una PCR cuantitativa de las regiones 5'-UTR y NS5b. El procedimiento propone cebadores y sondas de NS5b específicos para la detección de los subgenotipos 1a y 1b, y otros cebadores y sondas específicos de 5'-UTR para los genotipos y subgenotipos 2a, 2b, 3, 4, 5 y 6.

Todas estas metodologías aparecen divulgadas en el documento de Chevaliez S. et. al. "Hepatitis C virus (HCV) genotype 1 subtype identification in new HCV drug development and future clinical practice", PLoS ONE, 2009, vol. 4, pp. 1-9.

Tanaka et al. (PNAS 2002, 99(24), pp. 15584-15589) divulgan un análisis predictivo de la incidencia del carcinoma hepatocelular en los Estados Unidos, incluyendo la divulgación varios oligonucleótidos específicos del VHC.

El documento WO 2008/125119 trata de la manipulación de recombinantes intergenotípicos del virus de la hepatitis C con respecto a varios genes estructurales.

El documento WO 03/057915 describe la detección del genotipo 1 del VHC basada en ensayos de amplificación usando cebadores específicos de 5'-NCR.

Park et al. (J.Vir.Met. (2010), 163(2), pp. 269-275) divulgan los esfuerzos en el genotipado del VHC utilizando una micromatriz de ácido nucleico recientemente desarrollada.

Nakatani et al. (Virology Journal (2011), 8(1), p. 459) comparan el rendimiento del genotipado del VHC basado en la secuenciación parcial de la región NS5B frente a la 5'-UTR.

La región NS5B también es la diana analizada en la publicación de Sandres-Saune et al. (J.Vir.Met. (2003) 109(2), pp. 187-193), con el propósito de genotipado del VHC.

El documento US 2009/325145 enseña el genotipado del VHC basándose en la información de las secuencias tanto de la región 5'-UTR como de la región central.

Las metodologías se suelen comparar con el procedimiento de referencia, que es la secuenciación de Sanger del virus completo, o de al menos la región NS5B o la región 5'UTR-núcleo/E1 (Simmonds et al., *supra*). Aunque la secuenciación se considera el procedimiento de referencia, implica las desventajas de lecturas erróneas debido al hecho de que la secuencia resultante corresponde a un promedio de las secuencias en una muestra aislada. Por lo tanto, no se pueden detectar coinfecciones (infecciones por más de un subtipo del VHC al mismo tiempo). Además, un procedimiento de clonación y secuenciación para detectar coinfecciones implica tiempos analíticos prolongados y costes elevados, lo que hace que no sea práctico en kits comerciales.

Si bien todos los kits existentes basados en secuenciación semiautomática, PCR cuantitativa o hibridación inversa tienen una alta tasa de éxito en la determinación del VHC, también existen algunas limitaciones que hacen que no sean infalibles y que den lugar a algunas determinaciones erróneas.

En efecto, hay algunos subtipos que resulta difícil distinguir entre sí, por ejemplo, los subtipos 1a y 1b, y existen algunos positivos falsos y negativos falsos. Además, los kits mencionados anteriormente no permiten la determinación de infecciones múltiples (coinfecciones) en el caso de que el sujeto esté infectado por un genotipo o subtipo adicional simultáneo en una baja carga o proporción.

Sin embargo, también existe un nivel adicional de complejidad en las infecciones por el VHC que dificulta mucho más la correcta determinación del genotipo o subgenotipo (subtipo), en particular la posibilidad de una infección por un virus recombinante. La existencia de variantes recombinantes del VHC se debe al hecho de que en pacientes coinfectados por más de un genotipo o subtipo, se produce una recombinación entre todas estas variantes que da lugar a un genotipo o subtipo "híbrido", lo que dificulta la determinación de variantes infecciosas reales. Como se indicó anteriormente, la determinación de la variante real es de especial importancia para adaptar o prescribir mejor un tratamiento médico.

Es, por lo tanto, de gran interés desarrollar adicionalmente ensayos de genotipado mejorados para detectar las variantes específicas en pacientes infectados por el VHC para diagnosticar mejor a los sujetos, y para detectar con mayor sensibilidad mutaciones del virus asociadas a resistencia, con el objetivo de pronosticar aún más las respuestas a los tratamientos prescritos contra la infección por el VHC.

SUMARIO DE LA INVENCIÓN

La presente invención proporciona kits y procedimientos que los usan para el genotipado y subtipado de muestras que contienen el virus de la hepatitis C (VHC) humano. Los cebadores de los kits de la presente invención

proporcionan una sensibilidad y especificidad clínica robusta, y proporcionan también un alto rendimiento y un flujo de trabajo eficiente.

5 En un primer aspecto, la presente divulgación describe un conjunto de cebadores para la transcripción inversa (retrotranscripción) y amplificación de las regiones 5'-UTR y un fragmento del núcleo (5'-núcleo) del ARN del VHC de una muestra aislada, incluyendo dicha combinación al menos un oligonucleótido de fórmula (I)

(N)_m-Z (I), como cebador inverso,

10 en la que

Z es un oligonucleótido que consiste en la SEQ ID NO: 1 (5'-CTAGTCGCGCGCACACCCA-3'),

15 m es un número entero que varía de 0 a 25 nucleótidos, y

N es un nucleótido seleccionado de A, T, C, G,

20 y un cebador directo que puede hibridar con la parte 5'-UTR-Núcleo del VHC, en el que la combinación de cebadores puede generar un fragmento que comprende desde los nucleótidos 146 a 490 de la SEQ ID NO: 2.

25 La secuencia (N)_m-Z o fórmula (I), siendo Z como se define anteriormente, también se puede denominar en la presente solicitud SEQ ID NO: 15.

Se ha encontrado que la combinación de cebadores divulgados en el presente documento puede generar un fragmento idéntico al que comprende desde los nucleótidos 146 a 490 de la SEQ ID NO: 2 cuando se somete a prueba frente a un subtipo 1a del VHC.

30 La SEQ ID NO: 2 corresponde a la región que comprende desde el nucleótido 1 al 914 de la secuencia consenso de la cepa aislada H77 (número de acceso en GenBank AF009606, versión 1 de 18 de junio de 2009). Esta región abarca las regiones completas 5'-UTR y de núcleo del genoma del virus, y la numeración se establece de acuerdo con el sistema que comienza en 1 para el primer nucleótido de la 5'-UTR. Un sistema de numeración normalizado en profundidad para nucleótidos, proteínas y epítomos del VHC se expone en el documento de Kuiken et al., "A Comprehensive System for Consistent Numbering of HCV Sequences, Proteins and Epitopes", *Hepatology* - 2006, DOI 10.1002/hep.21377, pp. 1355-1361. La secuencia de la cepa aislada H77 corresponde a un gen de poliproteína del subtipo 1a del virus de la hepatitis C, cds completas, con ARN lineal de 9646 pb. Kuiken et al. (*supra*) también establecen el sistema de numeración para otras variantes del VHC ya divulgadas en el estado de la técnica, así como para las que se identifiquen en el futuro. En particular, se divulga que, aunque el sistema de numeración se basa en una secuencia de referencia del genotipo 1a, cuando se usan varias herramientas de localizadores de secuencias con los diferentes genotipos del VHC existentes, estas herramientas alinean las secuencias de todas ellas de forma inequívoca con respecto a la secuencia de referencia. A partir de esta referencia, por lo tanto, se puede concluir que todas las variantes del VHC incluirán una región 5'-UTR-Núcleo con algunas eliminaciones, inserciones, sustituciones cuando se comparen con la misma región en la referencia del genotipo 1a. Si dicha secuencia incluye inserciones o eliminaciones, los criterios de Kuiken et al. también contemplan cómo citar esta modificación con respecto a la secuencia de referencia (subtipo 1a) sin cambiar la enumeración básica.

45 El conjunto de cebadores divulgado en el presente documento usado para generar el fragmento del genoma del VHC que incluye la región definida como 5'-UTR-Núcleo también puede emplearse para amplificar esta región en cualquiera de los otros genotipos y subtipos del VHC distintos del genotipo 1a. Por lo tanto, con solo algunos cebadores seleccionados, se obtiene una amplificación de confianza de cualquiera de las variantes de una muestra de un sujeto. Dicha amplificación da lugar a un fragmento de la región del genoma vírico con una longitud que comprende al menos 344 posiciones de nucleótidos de 5'-UTR-Núcleo y los nucleótidos de los cebadores directo e inverso, en el que se debe entender que la posición del nucleótido incluye cualquier nucleótido o una eliminación o interrupción de un nucleótido con respecto a la región de referencia del genotipo 1a. Es decir, si se localiza una interrupción de un nucleótido (con respecto a la secuencia de referencia) los nucleótidos precedente y posterior no cambiarán la numeración recibida de acuerdo con la secuencia de referencia, y el fragmento generado será un nucleótido más corto, pero que puede alinearse con las posiciones de nucleótidos globales de la secuencia de referencia que no lleva dicha eliminación. Del mismo modo, si se inserta un nucleótido con respecto a la secuencia de referencia, los nucleótidos precedente y posterior no cambiarán la numeración y entonces el fragmento generado será más largo e incluirá, aparte de las posiciones de nucleótidos de la secuencia de referencia, las adicionales. Para un ejemplo de la numeración de inserciones y eliminaciones con respecto a la secuencia de referencia, véase Kuiken et al. (*supra*).

60 En otras palabras, aunque en la presente divulgación el fragmento generado se define con referencia a la SEQ ID NO: 2, que corresponde a una cepa aislada clasificada como subtipo 1a, si la variante del VHC en la muestra sometida a prueba es de un genotipo o subtipo diferente entre los genotipos, también se producirá amplificación y estará presente un fragmento que comprende al menos los nucleótidos correspondientes de 146 a 490 de la SEQ ID

NO: 2 en las variantes sometidas a prueba. Si dicha secuencia incluye inserciones o eliminaciones, con respecto a la secuencia de referencia (subtipo 1a), esto se detectará por las herramientas de alineación y la enumeración de la variante de secuencia puede realizarse siguiendo los criterios de Kuiken et al. si es necesario.

5 Otro aspecto descrito en el presente documento es un oligonucleótido de fórmula (I) como se define anteriormente.

Otro aspecto descrito en el presente documento es un conjunto de cebadores que comprende el oligonucleótido de fórmula (I) como se define anteriormente. Otro aspecto descrito en el presente documento es una composición que comprende el oligonucleótido de fórmula (I).

10 Un aspecto adicional descrito en el presente documento es un conjunto de cebadores que comprende al menos un oligonucleótido de fórmula (II).

(N)_m-F (II), como cebador inverso,

15 en la que

F es un oligonucleótido que consiste en la SEQ ID NO: 5 (5'-TTNGADGAGCADGATGTWATBAGCTC-3'), en la que

20 N significa adenina (A), citosina (C), guanina (G) o timina (T);

D significa G, A o T;

25 W significa A o T,

B significa G, T o C, y

m es un número entero que varía de 0 a 25 nucleótidos; y

30 un cebador directo que puede hibridar con la parte NS5B del VHC, en el que el conjunto de cebadores puede generar un fragmento que comprende desde los nucleótidos 653 a 1106 de la SEQ ID NO: 6.

35 Se ha encontrado que la combinación de cebadores de este aspecto descrita en el presente documento puede generar un fragmento idéntico al que comprende los nucleótidos 653 a 1106 de la SEQ ID NO: 6 cuando se somete a prueba frente a un subtipo 1a.

40 La SEQ ID NO: 6 corresponde a la región que comprende desde el nucleótido 7602 hasta el 9377 de la secuencia consenso de la cepa aislada H77 divulgada anteriormente (número de acceso en GenBank AF009606, versión 1 de 18 de junio de 2009). Esta región abarca toda la región NS5B del genoma del virus, y la numeración se establece de acuerdo con el sistema que comienza en 1 para el primer nucleótido de la 5'-UTR como se expuso anteriormente. Por lo tanto, con respecto a la secuencia consenso de la cepa aislada H77, el conjunto de cebadores puede generar un fragmento que comprende desde los nucleótidos 8254 a 8707 del consenso. Este fragmento generado, cuando se indica con respecto a la SEQ ID NO: 6 corresponde a un fragmento que comprende desde los nucleótidos 653 a 1106 de dicha SEQ ID NO: 6. Para la región NS5B, el fragmento generado se define con referencia a la SEQ ID NO: 6, que corresponde a una cepa aislada clasificada como subtipo 1a. Como se indica para la SEQ ID NO: 2 y considerando la nomenclatura establecida de la referencia de Kuiken et al. (*supra*), las otras variantes (genotipos o incluso subtipos entre genotipos) del VHC pueden incluir en la región NS5B algunas eliminaciones, inserciones, sustituciones cuando se comparan con la misma región en la referencia del genotipo 1a. Si dicha secuencia incluye inserciones o eliminaciones, los criterios de Kuiken et al. también contemplan cómo citar esta modificación con respecto a la secuencia de referencia (subtipo 1a) sin cambiar la enumeración básica.

55 El conjunto de cebadores descrito en el presente documento usado para generar el fragmento del genoma del VHC que incluye la región definida como región NS5B también se puede emplear para amplificar esta región en cualquiera de los otros genotipos y subtipos del VHC distintos del genotipo 1a.

60 En otras palabras, si la variante del VHC en la muestra sometida a prueba es diferente del genotipo 1a, también se producirá amplificación y estará presente un fragmento que comprende al menos los correspondientes nucleótidos 653 a 1106 de la SEQ ID NO: 6 en las variantes sometidas a prueba o, lo que es lo mismo, el fragmento correspondiente de los nucleótidos 8254 a 8707 de la secuencia consenso de la cepa aislada H77. Como antes, los cebadores amplificarán la secuencia correspondiente y si dicha secuencia incluye inserciones o eliminaciones, la propuesta de Kuiken et al. (*supra*) contempla cómo citar esta modificación con respecto a la secuencia de referencia (genotipo 1a) sin cambiar la enumeración básica. Por lo tanto, el conjunto de cebadores de acuerdo con la invención utilizado para generar el fragmento del genoma del VHC que incluye la región definida como NS5B da lugar a un fragmento de esta región del genoma del virus con una longitud que comprende al menos 354 posiciones de nucleótidos de esta región y los nucleótidos de los cebadores directo e inverso, en los que se debe entender que la

posición de los nucleótidos incluye cualquier nucleótido o una eliminación o interrupción de un nucleótido con respecto a la región de referencia del genotipo 1a.

5 Otro aspecto descrito en el presente documento es un oligonucleótido de fórmula (II) como se divulga anteriormente. Esta fórmula (II) también se puede definir como un oligonucleótido de la SEQ ID NO: 16, que incluye F como la SEQ ID NO: 5, N como un nucleótido seleccionado de A, T, C, G, y en la que m es de 0 a 25.

10 Otro aspecto descrito en el presente documento es un conjunto de cebadores que comprende el oligonucleótido de fórmula (II) como se define anteriormente. Otro aspecto descrito en el presente documento es una composición que comprende el oligonucleótido de fórmula (II).

15 Estos conjuntos de cebadores descritos en el presente documento, usados en combinación, permiten transcribir inversamente (transcripción inversa) y amplificar aquellas regiones clave del ARN del VHC, incluidas las mutaciones existentes entre los genotipos y subtipos del virus. Además, estos cebadores son lo suficientemente específicos de estas regiones informativas clave, por lo que dan lugar a amplicones de alta pureza que pueden secuenciarse con un alto nivel de confianza para determinar la variante o variantes reales del VHC (genotipo, subtipo) en una muestra.

20 Otro aspecto descrito en el presente documento es un oligonucleótido que consiste en la SEQ ID NO: 7 (5'-CNTAYGAYACCMGNTGYTTTGACTC-3'), en la que N significa adenina (A), citosina (C), guanina (G) o timina (T); Y significa T o C; y M significa A o C.

25 Otro aspecto más descrito en el presente documento es un conjunto de cebadores que comprende el oligonucleótido que consiste en la SEQ ID NO: 7 como cebador directo, y un cebador inverso que puede hibridar con la parte NS5B del VHC, en el que el conjunto de cebadores puede generar un fragmento que comprende los nucleótidos 653 a 1040 de la SEQ ID NO: 6.

30 Este fragmento de los nucleótidos 653 a 1040 de la SEQ ID NO: 6 es el fragmento equivalente (o correspondiente), indicado con respecto a la secuencia consenso de la cepa aislada H77, que comprende desde los nucleótidos 8254 a 8641 del consenso. Como se expuso anteriormente, la numeración de los nucleótidos se establece de acuerdo con el sistema que comienza en 1 para el primer nucleótido de la 5'-UTR de la secuencia consenso H77 (número de acceso en GenBank AF009606).

35 Otro aspecto más descrito en el presente documento es un oligonucleótido que consiste en la SEQ ID NO: 8 (5'-GARTAYCTGGTCATAGCNTCCGTGAA-3'), en la que R significa adenina (A) o guanina (G); Y significa timina (T) o citosina (C); y N significa adenina (A), citosina (C), guanina (G) o timina (T).

40 Este conjunto de cebadores que incluyen el oligonucleótido que consiste en la SEQ ID NO: 7 como cebador directo, y que abarca varios codones de la parte NS5B del VHC para generar un fragmento que comprende los nucleótidos 653 a 1040 de la SEQ ID NO: 6, es especialmente útil para mejorar la pureza de cualquier fragmento previamente transcrito inversamente y amplificado de esta misma región. El amplicón resultante todavía contiene las regiones clave que incluyen las mutaciones entre genotipos y subtipos del VHC. Por lo tanto, proporcionan una muestra altamente pura con el objetivo de seguir su secuenciación y determinar la variante del VHC.

45 Como se ilustrará en los ejemplos a continuación, todos estos conjuntos de cebadores, utilizados en combinación en un procedimiento *in vitro* para determinar el genotipo o subtipo de VHC en una muestra, permiten mejores puntuaciones en el porcentaje de identificaciones correctas en relación con los otros kits y procedimientos de la técnica anterior. Estos resultados son incluso mejores que los logrados con los procedimientos de secuenciación de referencia en los casos especiales de muestras analizadas coinfectadas.

50 También se divulga en el presente documento un kit para detectar una o más variantes de secuencia del VHC que comprende al menos uno de los conjuntos de cebadores como se define anteriormente.

55 También se divulga en el presente documento el uso de cualquiera de los conjuntos de cebadores mencionados anteriormente solos o en combinación para detectar una o más variantes de secuencia del VHC.

Un aspecto adicional de la invención proporciona un procedimiento *in vitro* para determinar en una muestra aislada de un sujeto la presencia de una o más variantes de secuencia del virus de la hepatitis C (VHC) humano que comprende:

60 i) transcribir inversamente y amplificar moléculas de ARN del VHC mediante reacción en cadena de la polimerasa con transcripción inversa (RT-PCR) utilizando el primer conjunto de cebadores como se define anteriormente en el presente documento, y el segundo conjunto de cebadores como se define anteriormente en el presente documento,

65 ii) amplificar adicionalmente los fragmentos generados en la etapa i) mediante PCR usando el primer conjunto de cebadores como se define anteriormente en el presente documento; y el tercer conjunto de cebadores como se define anteriormente en el presente documento, y

iii) secuenciar los fragmentos amplificados obtenidos en la etapa ii).

BREVE DESCRIPCIÓN DE LOS DIBUJOS

- 5 La figura 1 es un diagrama de árbol filogenético de los diferentes subtipos del VHC que muestra el emplazamiento de una de las secuencias amplificadas detectadas en una muestra de prueba de un paciente (paciente P43, Px10.66) infectado por el subtipo 1a.
- 10 La figura 2 es también otro diagrama de árbol filogenético, en el que se emplaza otra de las secuencias amplificadas (paciente P43, Px1.1506rd) de la muestra sometida a prueba (P43) del mismo paciente de la figura 1. En este caso, se muestra la identificación de la infección por el subtipo 4d.

DESCRIPCIÓN DETALLADA DE LA INVENCIÓN

- 15 A menos que se definan de otro modo, todos los términos técnicos y científicos usados en el presente documento tienen, en general, el mismo significado que entiende comúnmente un experto en la técnica a la que pertenece esta invención. En general, la nomenclatura usada en el presente documento y los procedimientos de laboratorio en cultivo celular, genética molecular, química orgánica y química e hibridación de ácidos nucleicos descritos a
- 20 continuación son los conocidos bien y comúnmente empleados en la técnica.

- "Secuencia variante del VHC" se refiere al clado o genotipo (incluidos los subtipos en el genotipo) del VHC específico, que se clasifica de acuerdo con un patrón conocido que enumera todas las variantes identificadas del virus. Un "genotipo" se define como la composición genética de una célula, un organismo (incluidos virus) o un individuo (es decir, la composición alélica específica del individuo) habitualmente con referencia a un carácter específico en consideración. En la mayoría de los casos, una secuencia variante y un genotipo identificado son términos intercambiables. Por otro lado, una secuencia variante puede ser el resultado de la recombinación genética de dos genotipos del VHC identificados. En este último caso, la secuencia variante del VHC no es equivalente a un genotipo o clado, sino a mezclas de genotipos. Los genotipos del VHC difieren entre sí en un 31 % a un 33 % a nivel de nucleótidos. Un "subtipo" o "subgenotipo" (usados de forma intercambiable en el presente documento) se define como una variante de secuencia a niveles de nucleótidos en un genotipo específico. Los subtipos del VHC difieren entre sí en un 20 % a un 25 % a nivel de nucleótidos, aunque pertenecen al mismo genotipo.

- Como se usa en el presente documento, el término "genotipado" o "determinación de una secuencia variante del virus" es el procedimiento de determinación de las diferencias en la composición genética (o genotipo) de un virus individual mediante el examen de la secuencia de ADN o ARN del virus utilizando ensayos biológicos y comparándola con otra secuencia de virus o una secuencia de referencia. Tradicionalmente, el genotipado es el uso de secuencias de ADN o ARN para definir poblaciones biológicas mediante el uso de herramientas moleculares. Por lo general, no implica la definición de los genes de un individuo.

- Como se usa en el presente documento, el término "amplicón" se refiere a un producto de una reacción de amplificación. Un ejemplo de un amplicón es un ADN o un producto de ARN (habitualmente un segmento de un gen, ADN o ARN) producido como resultado de PCR, PCR en tiempo real, PCR con transcripción inversa (RT-PCR), PCR en tiempo real competitiva, reacción en cadena de la ligasa (LCR), gap-LCR, amplificación por desplazamiento de hebra (SDA), amplificación basada en secuencia de ácidos nucleicos (NASBA), amplificación mediada por transcripción (TMA) o similar. Como se usan en el presente documento, las frases "amplificación", "procedimiento de amplificación" o "reacción de amplificación" se usan de forma intercambiable y se refieren a un procedimiento que aumenta la representación de una población de secuencias específicas (tales como una secuencia diana o un ácido nucleico diana) de ácidos nucleicos (todos los tipos de ADN o ARN) en una muestra de prueba. Los ejemplos de procedimientos de amplificación que se pueden usar en la presente invención incluyen, pero sin limitación, PCR, PCR en tiempo real, PCR con transcripción inversa (RT-PCR), PCR en tiempo real competitiva, reacción en cadena de la ligasa (LCR), gap-LCR, amplificación por desplazamiento de hebra (SDA), amplificación basada en secuencia de ácidos nucleicos (NASBA), amplificación mediada por transcripción (TMA) y similares, que son todos ellos conocidos por los expertos en la técnica.

- Como se usa en el presente documento, el término "hibridación" se refiere a la formación de complejos entre secuencias de ácido nucleico que son suficientemente complementarias para formar complejos a través del apareamiento de bases de Watson y Crick o el apareamiento de bases no canónico. Por ejemplo, cuando un cebador "hibrida" con una secuencia diana (plantilla), dichos complejos (o híbridos) son suficientemente estables para ejercer la función de cebado requerida por, por ejemplo, la ADN-polimerasa, para iniciar la síntesis del ADN. Un experto en la técnica apreciará que las secuencias de hibridación no necesitan tener complementariedades perfectas para proporcionar híbridos estables. En muchas situaciones, se formarán híbridos estables cuando haya errores de apareamiento en menos de aproximadamente un 10 % de las bases. Por consiguiente, como se usa en el presente documento, el término "complementario" se refiere a un oligonucleótido que forma un dúplex estable con su complemento en condiciones de ensayo, en general cuando hay aproximadamente un 80 %, aproximadamente un 81 %, aproximadamente un 82 %, aproximadamente un 83 %, aproximadamente un 84 %, aproximadamente un

85 %, aproximadamente un 86 %, aproximadamente un 87 %, aproximadamente un 88 %, aproximadamente un 89 %, aproximadamente un 90 %, aproximadamente un 91 %, aproximadamente un 92 %, aproximadamente un 93 %, aproximadamente un 94 %, aproximadamente un 95 %, aproximadamente un 96 %, aproximadamente un 97 %, aproximadamente un 98 % o aproximadamente un 99 % de apareamiento de bases de Watson y Crick. Los expertos en la técnica entienden cómo estimar y ajustar la rigurosidad de las condiciones de hibridación de modo que las secuencias que tengan al menos un nivel deseado de complementariedades se hibriden de forma estable, mientras que las que tengan menores complementariedades no lo hagan.

El término "cebador" se refiere a un oligonucleótido que puede actuar como un punto de iniciación de la síntesis de un producto de extensión de cebador que es una cadena complementaria de ácido nucleico (todos los tipos de ADN o ARN), cuando se coloca en condiciones de amplificación adecuadas (por ejemplo, tampón, sal, temperatura y pH) en presencia de nucleótidos y un agente para la polimerización de ácidos nucleicos (por ejemplo, una polimerasa dependiente de ADN o dependiente de ARN). El cebador puede ser monocatenario o bicatenario. Si es bicatenario, el cebador se puede tratar primero (por ejemplo, desnaturalizarse) para permitir la separación de sus cadenas antes de usarlo para preparar productos de extensión. Dicha etapa de desnaturalización se realiza típicamente usando calor, pero de forma alternativa puede llevarse a cabo usando álcali, seguido de neutralización. Los cebadores de la presente invención pueden tener una longitud de aproximadamente 15 a aproximadamente 50 nucleótidos de longitud, preferentemente de aproximadamente 20 a aproximadamente 40 nucleótidos de longitud, lo más preferentemente, de aproximadamente 22 a aproximadamente 30 nucleótidos de longitud. Los cebadores de la presente invención pueden contener nucleótidos adicionales además de los descritos con más detalle en el presente documento. La frase "cebador directo" se refiere a un cebador que hibrida (o se aparea) con la secuencia diana (por ejemplo, cadena plantilla). La frase "cebador inverso" se refiere a un cebador que hibrida (o se aparea) con la cadena complementaria de la secuencia diana. El cebador directo hibrida con la secuencia diana 5' con respecto al cebador inverso.

Como se usa en el presente documento, la expresión "conjunto de cebadores" se refiere a dos o más cebadores que juntos pueden cebar la amplificación de una secuencia diana o ácido nucleico diana de interés (por ejemplo, una secuencia diana dentro del VHC). En ciertos modos de realización, el término "conjunto de cebadores" se refiere a un par de cebadores que incluye un cebador 5' (en dirección 5') (o cebador directo) que hibrida con el extremo 5' complementario de la secuencia diana o ácido nucleico diana que se va a amplificar y un cebador 3' (en dirección 3') (o cebador inverso) que hibrida con el complemento de la secuencia diana o ácido nucleico diana que se va a amplificar. Dichos conjuntos de cebadores o pares de cebadores son particularmente útiles en reacciones de amplificación por PCR.

Un "oligonucleótido" se debe entender como una molécula de ARN o ADN monocatenario corta (habitualmente de 2 a 20 nucleótidos) que tiene una amplia variedad de aplicaciones en pruebas genéticas, investigación y análisis forense. De hecho, un cebador es un oligonucleótido. Frecuentemente preparados en el laboratorio por síntesis química en fase sólida, estos pequeños trozos de ácidos nucleicos pueden fabricarse con cualquier secuencia especificada por el usuario y, por lo tanto, son vitales para la síntesis de genes artificiales, reacción en cadena de la polimerasa (PCR), secuenciación de ADN, construcción de bibliotecas y como sondas moleculares. En la naturaleza, los oligonucleótidos se encuentran habitualmente como pequeñas moléculas de ARN que funcionan en la regulación de la expresión génica (por ejemplo, microARN), o son productos intermedios de degradación derivados de la descomposición de moléculas de ácido nucleico más grandes. Los oligonucleótidos se unen fácilmente, de una manera específica de la secuencia, a sus respectivos oligonucleótidos complementarios, ADN o ARN para formar dúplex o, con menos frecuencia, híbridos de un orden superior.

Si no se expresa lo contrario, cuando en la presente invención se indica una secuencia de un cebador o de un fragmento del VHC, el extremo 5' de la secuencia corresponde al primer nucleótido indicado en la secuencia, de acuerdo con la nomenclatura estándar aceptada en la que las secuencias de nucleótidos se expresan en el sentido de 5' a 3' (siguiendo los criterios de Kuiken et al. como se analiza en profundidad anteriormente).

A veces se usan cebadores degenerados. Estos son, en realidad, mezclas de cebadores similares, pero no idénticos. Pueden ser útiles si la región genética con la que hibrida el cebador muestra una variación en posiciones de nucleótidos particulares cuando se comparan diferentes cepas aisladas (genotipos o subtipos).

Las frases "secuencia diana" y "ácido nucleico diana" se usan de forma intercambiable en el presente documento y se refieren a la secuencia cuya presencia o ausencia se desea detectar. En el contexto de la presente invención, una secuencia diana incluye preferentemente una secuencia de ácido nucleico con la que formarán complejo uno o más cebadores. La secuencia diana también puede incluir una región de hibridación con sonda con la que una sonda formará un híbrido estable en condiciones de amplificación apropiadas. Como reconocerá un experto en la técnica, una secuencia diana puede ser monocatenaria o bicatenaria.

Como se usa en el presente documento, el término "muestra de prueba" en general se refiere a un material biológico que se somete a prueba y/o que se sospecha que contiene un analito de interés, tal como una secuencia de nucleótidos del VHC. La muestra de prueba puede derivarse de cualquier fuente biológica, tal como un hisopo o cepillo cervical, vaginal o anal, o un líquido fisiológico, incluyendo, pero sin limitación, sangre entera, suero, plasma,

líquido intersticial, saliva, líquido ocular del cristalino, líquido cefalorraquídeo, sudor, orina, leche, líquido ascítico, moco, líquido nasal, esputo, líquido sinovial, líquido peritoneal, secreción vaginal, menstruación, líquido amniótico, semen, y así sucesivamente. La muestra de prueba se puede usar directamente como se obtiene de la fuente biológica o después de un pretratamiento para modificar el carácter de la muestra. Por ejemplo, dicho pretratamiento puede incluir preparar plasma a partir de sangre, diluir líquidos viscosos, y así sucesivamente. Los procedimientos de pretratamiento también pueden incluir filtración, precipitación, dilución, destilación, mezcla, concentración, liofilización, inactivación de componentes interferentes, adición de reactivos, lisis, etc. Además, también puede ser beneficioso modificar una muestra de prueba sólida para formar un medio líquido o para liberar el analito. Preferentemente, la muestra puede ser suero o plasma.

El principal problema al diseñar cebadores para la transcripción inversa y la amplificación del VHC para una secuenciación posterior es el alto número de genotipos y subtipos identificados para el VHC. Esto hace que la selección de la región genómica de interés sea una tarea engorrosa.

En principio, los cebadores deben diseñarse en regiones altamente conservadas de los virus. Esto es esencial para garantizar un rendimiento fiable de los cebadores diseñados para amplificar la región genómica de interés sin un riesgo sustancial de fracaso. El fracaso de la amplificación se debe principalmente a la mutación en la región diana que hace que el cebador no pueda unirse. Esta tarea es aún más difícil cuando se diseñan cebadores para el genotipado del VHC, ya que los cebadores deben diseñarse de modo que el mismo conjunto de cebadores se pueda usar para secuenciar la región diana para múltiples clados.

Sorprendentemente, los oligonucleótidos divulgados en el presente documento se han diseñado de modo que puedan usarse como cebadores que pueden transcribir inversamente y amplificar esas regiones clave de la secuencia del ARN del VHC, incluidas las mutaciones existentes entre los genotipos y subtipos del virus. Además, estos cebadores son lo suficientemente específicos de estas regiones clave, por lo que dan lugar a amplicones de alta pureza que se pueden secuenciar con un alto nivel de confianza para determinar la variante o variantes reales del VHC (genotipo, subtipo) en una muestra.

En algunos casos del oligonucleótido de fórmula (I), m es 20. En otros casos, m es 10 y, en otros casos más, m es 5. En otros casos m se selecciona de 1, 2, 3, 4, 6, 7, 8, 9, 11, 12, 13, 14, 15, 16, 17, 18, 19, 21, 22, 23, 24 y 25.

En un modo de realización preferente, el oligonucleótido de fórmula (I) consiste en la SEQ ID NO: 1, por lo tanto, es un oligonucleótido de fórmula (I) en la que m es 0.

Cuando el oligonucleótido de fórmula (I) es un cebador inverso que forma parte de un conjunto de cebadores de acuerdo con la invención, el cebador directo que puede hibridar con la parte 5'-UTR-Núcleo del VHC consiste en una secuencia seleccionada de la SEQ ID NO: 3 y la SEQ ID NO: 4.

En algunos casos del oligonucleótido de fórmula (II), m es 20. En otro modo de realización m es 10, y en otro modo de realización más m es 5. En otros modos de realización m se selecciona de 1, 2, 3, 4, 6, 7, 8, 9, 11, 12, 13, 14, 15, 16, 17, 18, 19, 21, 22, 23, 24 y 25.

También se divulga en el presente documento un oligonucleótido de fórmula (II) en la que F es un oligonucleótido que consiste en la SEQ ID NO: 5 (5'- TTNGADGAGCADGATGTWATBAGCTC-3'), es decir, m es 0.

De hecho, la SEQ ID NO: 5 incluye 216 secuencias diferentes que consideran las múltiples combinaciones de los valores N, D, W y B. Las secuencias preferentes se seleccionan del grupo que consiste en la SEQ ID NO: 36 a la SEQ ID NO: 53. Es decir, TTAGATGAGCATGATGTAATGAGCTC (SEQ ID NO: 36),

TTAGATGAGCATGATGTAATTAGCTC (SEQ ID NO: 37),

TTAGATGAGCATGATGTAATCAGCTC (SEQ ID NO: 38),

TTAGATGAGCATGATGTTATGAGCTC (SEQ ID NO: 39),

TTAGATGAGCATGATGTTATTAGCTC (SEQ ID NO: 40),

TTAGATGAGCATGATGTTATCAGCTC (SEQ ID NO: 41),

TTAGATGAGCAGGATGTAATGAGCTC (SEQ ID NO: 42),

TTAGATGAGCAGGATGTAATTAGCTC (SEQ ID NO: 43),

TTAGATGAGCAGGATGTAATCAGCTC (SEQ ID NO: 44),

TTAGATGAGCAGGATGTTATGAGCTC (SEQ ID NO: 45),

TTAGATGAGCAGGATGTTATTAGCTC (SEQ ID NO: 46),

TTAGATGAGCAGGATGTTATCAGCTC (SEQ ID NO: 47),

TTAGATGAGCAAGATGTAATGAGCTC (SEQ ID NO: 48),

TTAGATGAGCAAGATGTAATTAGCTC (SEQ ID NO: 49),

TTAGATGAGCAAGATGTAATCAGCTC (SEQ ID NO: 50),

TTAGATGAGCAAGATGTTATGAGCTC (SEQ ID NO: 51),

TTAGATGAGCAAGATGTTATTAGCTC (SEQ ID NO: 52) y
TTAGATGAGCAAGATGTTATCAGCTC (SEQ ID NO: 53)

5 En un modo de realización preferente, cuando el oligonucleótido de fórmula (II) es un cebador inverso que forma parte de un conjunto de cebadores, el cebador directo que puede hibridar con la parte NS5B del VHC consiste en la SEQ ID NO: 7 (5'-CNTAYGAYACCMGNTGYTTTGA^{CTC}-3'), en la que N significa adenina (A), citosina (C), guanina (G) o timina (T);

10 Y significa T o C; y M significa A o C.

La SEQ ID NO: 7 incluye 256 secuencias diferentes que consideran las múltiples combinaciones de los valores N, Y y M. Las secuencias preferentes se seleccionan del grupo que consiste en la SEQ ID NO: 54 a la SEQ ID NO: 69. Es decir, CATATGATACCAGATGCTTTGACTC (SEQ ID NO: 54),

15 CATATGATACCAGATGTTTTGACTC (SEQ ID NO: 55),
CATATGATACCAGTTGCTTTGACTC (SEQ ID NO: 56),
CATATGATACCAGTTGTTTTGACTC (SEQ ID NO: 57),
CATATGATACCAGGTGCTTTGACTC (SEQ ID NO: 58),
CATATGATACCAGGTGTTTTGACTC (SEQ ID NO: 59),
CATATGATACCAGCTGCTTTGACTC (SEQ ID NO: 60),
20 CATATGATACCAGCTGTTTTGACTC (SEQ ID NO: 61),
CATATGATACCCGATGCTTTGACTC (SEQ ID NO: 62),
CATATGATACCCGATGTTTTGACTC (SEQ ID NO: 63),
CATATGATACCCGTTGCTTTGACTC (SEQ ID NO: 64),
CATATGATACCCGTTGTTTTGACTC (SEQ ID NO: 65),
25 CATATGATACCCGGTGCTTTGACTC (SEQ ID NO: 66),
CATATGATACCCGGTGTTTTGACTC (SEQ ID NO: 67),
CATATGATACCCGCTGCTTTGACTC (SEQ ID NO: 68),
CATATGATACCCGCTGTTTTGACTC (SEQ ID NO: 69)

30 Cuando un oligonucleótido de fórmula (II) es un cebador inverso, las secuencias preferentes son aquellas seleccionadas del grupo que consiste en la SEQ ID NO: 36 a la SEQ ID NO: 53 y mezclas de las mismas. En un conjunto de cebadores que incluye oligonucleótidos degenerados, se emplean mezclas de todos los cebadores para asegurar la amplificación de todas las variantes del VHC en una muestra.

35 De la misma manera, cuando se utiliza como cebador directo un oligonucleótido degenerado que consiste en la SEQ ID NO: 7, las secuencias preferentes se seleccionan del grupo que consiste en la SEQ ID NO: 54 a la SEQ ID NO: 69, y mezclas de las mismas.

40 La presente divulgación también proporciona un conjunto de cebadores especialmente diseñados para realizar amplificaciones hemianidadas de regiones (fragmentos) previamente transcritas inversamente y amplificadas. Este conjunto de cebadores incluye un cebador u oligonucleótido que consiste en la SEQ ID NO: 7 como cebador directo, y un cebador inverso que puede hibridar con la parte NS5B del VHC, generando el conjunto un fragmento que comprende del nucleótido 653 al 1040 de la SEQ ID NO: 6.

45 En un modo de realización preferente, dicho cebador inverso en el conjunto consiste en un oligonucleótido de la SEQ ID NO: 8 (Bo8641).

50 De hecho, el oligonucleótido de la SEQ ID NO: 8 representa un grupo de 16 secuencias, cuando se consideran las múltiples combinaciones de los valores R, Y y N. Las secuencias preferentes se seleccionan del grupo que consiste en la SEQ ID NO: 70 a la SEQ ID NO: 73. Es decir, GAATATCTGGTCATAGCATCCGTGAA (SEQ ID NO: 70), GAATATCTGGTCATAGCTTCCGTGAA (SEQ ID NO: 71), GAATATCTGGTCATAGCGTCCGTGAA (SEQ ID NO: 72) y GAATATCTGGTCATAGCCTCCGTGAA (SEQ ID NO: 73).

55 Como se expuso anteriormente, cuando se utiliza como cebador directo un oligonucleótido degenerado que consiste en la SEQ ID NO: 7, las secuencias preferentes se seleccionan del grupo que consiste en la SEQ ID NO: 54 a la SEQ ID NO: 69, y mezclas de las mismas.

60 De la misma manera, cuando se utiliza como cebador inverso un oligonucleótido degenerado que consiste en la SEQ ID NO: 8, las secuencias preferentes se seleccionan del grupo que consiste en la SEQ ID NO: 70 a la SEQ ID NO: 73 y mezclas de las mismas.

Con solo estos tres tipos de conjuntos de cebadores, empleados en combinación en un procedimiento para

determinar la presencia de una o más variantes de secuencia del VHC en una muestra, es posible detectar cualquiera de los genotipos o subtipos del VHC, aunque una variante específica esté presente a niveles muy bajos en dicha muestra.

5 Todos estos conjuntos de cebadores pueden formar parte de un kit para genotipar el VHC, en cualquiera de las combinaciones posibles de los conjuntos, o se pueden concebir como reactivos independientes útiles en un procedimiento para genotipar el VHC. Es decir, un kit puede contener solo uno del conjunto de cebadores, o dos, o tres.

10 El kit puede comprender adicionalmente los medios (aditivos, disolventes) para realizar la transcripción inversa y/o amplificación de un fragmento de gen deseado, incluyendo la solución tampón, las enzimas (polimerasas), mezclas de nucleótidos (dNTPs, ddNTPs), aditivos para mejorar la visualización de un fragmento amplificado en un gel de electroforesis, tales como sondas (específicas o no). Algunos de los componentes adecuados para formar parte de los kits de la invención se proporcionan en los ejemplos. Además, los conjuntos de cebadores de acuerdo con la invención también pueden ser integrantes o reactivos de diferentes kits comerciales. Los ejemplos de kits comerciales incluyen el kit Transcriptor one-step RT-PCR de Roche, o el sistema de PCR FastStart High Fidelity, dNTPack (Roche).

20 La presente invención también proporciona un procedimiento *in vitro* para determinar la variante de secuencia del VHC en una muestra aislada de un sujeto, de forma que en dicho procedimiento los conjuntos de cebadores divulgados anteriormente se usan en combinación para genotipar aquellas regiones críticas que incluyen las mutaciones significativas que permiten la distinción entre genotipos y subtipos en un genotipo.

25 En un modo de realización preferente de este procedimiento, la etapa i) de transcripción inversa y amplificación de moléculas de ARN del VHC se realiza usando el siguiente conjunto de cebadores:

a) un cebador inverso de fórmula (I) y un cebador directo que consiste en la SEQ ID NO: 3 (UTR45), y

30 b) un cebador inverso de fórmula (II) y un cebador directo que consiste en la SEQ ID NO: 7 (Bo8245).

En un modo de realización más preferente, el cebador inverso en a) consiste en la SEQ ID NO: 1, y el cebador inverso en b) consiste en la SEQ ID NO: 5. Los cebadores inversos preferentes se seleccionan del grupo que consiste en la SEQ ID NO: 36 a la SEQ ID NO: 53 y mezclas de las mismas.

35 En otro modo de realización del procedimiento de acuerdo con la invención, la etapa ii) se realiza usando los siguientes conjuntos de cebadores:

a) un cebador inverso de fórmula (I) y un cebador directo que consiste en la SEQ ID NO: 4, y

40 b) un cebador directo que consiste en la SEQ ID NO: 7 y un cebador inverso que consiste en la SEQ ID NO: 8.

45 En este procedimiento, los cebadores directos preferentes genéricamente definidos como SEQ ID NO: 7 se seleccionan del grupo que consiste en la SEQ ID NO: 54 a la SEQ ID NO: 69, y mezclas de las mismas. También los cebadores inversos preferentes genéricamente definidos como SEQ ID NO: 8 se seleccionan del grupo que consiste en la SEQ ID NO: 70 a la SEQ ID NO: 73, y mezclas de las mismas.

50 Una vez que los fragmentos de interés se han amplificado, se someten además a una etapa de secuenciación. Esta etapa puede llevarse a cabo mediante cualquiera de las tecnologías de secuenciación conocidas, tales como el procedimiento de Sanger o el procedimiento de pirosecuenciación.

55 Si se elige el procedimiento de Sanger, el par de cebadores definido por el cebador inverso de fórmula (I), que preferentemente consiste en la SEQ ID NO: 1, con un cebador directo de la SEQ ID NO: 4, se puede usar para la secuenciación de la zona (región) 5'-UTR-núcleo de las variantes del VHC. En paralelo, el par de cebadores definidos por el cebador directo que consiste en la SEQ ID NO: 7 con el cebador inverso que consiste en la SEQ ID NO: 8 se puede emplear para secuenciar la zona (región) NS5B de las variantes del VHC.

60 Si se emplea pirosecuenciación, es muy preferente cualquier tecnología de secuenciación de lectura larga de alto rendimiento conocida en la técnica que permita la secuenciación de muchas moléculas de ácido nucleico diana en paralelo generando miles de lecturas clonales. En un modo de realización preferente de la presente invención, la tecnología de secuenciación utilizada es la denominada secuenciación 454, descrita en detalle en los ejemplos a continuación.

65 La secuenciación 454 es una tecnología de secuenciación paralela masiva para generar miles de lecturas clonales de uno o múltiples sujetos en un único experimento de secuenciación. Por lo tanto, se puede lograr la evaluación simultánea de múltiples variantes del VHC con alta sensibilidad.

Sorprendentemente, los cebadores divulgados en el presente documento permiten la secuenciación de las zonas críticas completas del VHC utilizando tan solo cuatro conjuntos de cebadores, definidos en general en un grupo de tres tipos de conjuntos.

5 En un modo de realización preferente, cuando se realiza el procedimiento de la invención mediante secuenciación 454, los cebadores de la invención usados en la PCR hemianidada (etapa ii) comprenden en su sección 5' una secuencia de oligonucleótidos universal correspondiente a un fragmento del bacteriófago M13. La secuencia de fragmentos del bacteriófago M13 es diferente para los cebadores directo e inverso de la amplificación hemianidada, siendo GTTGTAAAACGACGGCCAGT (SEQ ID NO: 9) para cebadores directos y
10 CACAGGAAACAGCTATGACC (SEQ ID NO: 10) para cebadores inversos.

Si la última etapa (iii) de la secuenciación se va a realizar utilizando la tecnología de secuenciación de alto rendimiento 454, el procedimiento puede incluir opcionalmente la realización de una nueva reacción de PCR con los cebadores universales (propiedad de Roche) con las secuencias
15 CGTATCGCCTCCCTCGCGCCAGTTGTA AAAACGACGGCCAGT (SEQ ID NO: 11) como cebador directo y CTATGCGCCTTGCCAGCCCGCCACAGGAAACAGCTATGACC (SEQ ID NO: 12) como cebador inverso. Estos cebadores universales permiten amplificar los fragmentos para secuenciarlos adicionalmente para que los amplicones finalmente incluyan en la sección 5' los adaptadores (oligos) CGTATCGCCTCCCTCGCGCCA (SEQ ID NO: 13) y CTATGCGCCTTGCCAGCCCGC (SEQ ID NO: 14)
20 necesarios para la pirosecuenciación o cualquier otro tipo de secuenciación universal. De hecho, los adaptadores en la sección 5' corresponden a secuencias que son complementarias a un oligonucleótido inmovilizado en una microesfera. En dicho caso, los adaptadores en la sección 5' permiten que los amplicones se adhieran a las microesferas para su secuenciación posterior.

25 La tecnología de secuenciación de alto rendimiento 454 también se podría realizar con los cebadores de la invención usados en la etapa ii) de amplificación hemianidada, siempre que los cebadores incluyan también en su sección 5' los adaptadores CGTATCGCCTCCCTCGCGCCA (SEQ ID NO: 13) para cebadores directos y CTATGCGCCTTGCCAGCCCGC (SEQ ID NO: 14) para cebadores inversos.

30 La utilidad de la tecnología 454 en la secuenciación del VHC parece ampliamente divulgada en el documento de Lauck M., et al. "Analysis of hepatitis C virus intra-host diversity across the coding region by ultra-deep sequencing". J. Virol. - 2012, Epub (publicación en línea, antes de la impresión). Esta referencia establece la utilidad de la tecnología de Roche en la secuenciación de los fragmentos del genoma del VHC, y muestra la profundidad a la que la tecnología puede detectar mutaciones infrecuentes de un subtipo específico, a saber, el 1a del VHC. Los cebadores utilizados para realizar la transcripción inversa y la amplificación son los del kit Superscript III High-Fidelity One-Step RT-PCR (Invitrogen, Life Technologies, Carlsbad, CA). A continuación, los amplicones se someten a diferentes procedimientos para prepararlos para la secuenciación 454.
35

40 Utilizando cualquiera de las tecnologías de secuenciación de lectura larga de alto rendimiento conocidas en la técnica, el procedimiento de la invención se puede usar para el genotipado de las variantes de secuencia del VHC derivadas de al menos una muestra obtenida de un único paciente en un modo muy rápido, así como las variantes de secuencia del VHC derivadas de al menos una muestra obtenida de varios pacientes sin retraso en el tiempo de análisis y diagnóstico.

45 En cualquiera de los modos de realización del procedimiento de la invención, se puede añadir una etapa iv) adicional de correlación de las variantes de secuencia detectadas con un patrón que asocia la variante detectada con la resistencia a fármacos del VHC.

50 Esta etapa adicional representa una herramienta interesante para los médicos que pueden prescribir rápidamente el tratamiento o la pauta médica correcta para un paciente determinado.

El procedimiento y el kit de la invención implican la gran ventaja de que se pueden detectar coinfecciones con una carga baja de una de las variantes del VHC en una única muestra, de forma que se determine el tratamiento más apropiado para el paciente. El procedimiento y los cebadores también permiten la detección en una muestra de infecciones por el VHC recombinante, lo que también da lugar a la mejor opción de tratamiento.
55

A lo largo de la descripción y las reivindicaciones, la palabra "comprende" y las variaciones de la palabra no pretenden excluir otras características técnicas, aditivos, componentes o etapas. Además, la palabra "comprende" abarca el caso de "que consiste en". Objetos, ventajas y características adicionales de la invención serán evidentes para los expertos en la técnica tras el examen de la descripción o pueden aprenderse mediante la práctica de la invención. Los siguientes ejemplos se proporcionan a modo de ilustración, y no pretenden ser limitantes de la presente invención. Además, la presente invención cubre todas las posibles combinaciones de modos de realización particulares y preferentes descritos en el presente documento.
60

EJEMPLOS

Los siguientes ejemplos permiten poner de manifiesto la utilidad y las ventajas del procedimiento *in vitro* de acuerdo con la invención, en el que se emplean oligonucleótidos específicos de la invención que actúan como cebadores.

5

Ejemplo 1. Resultados comparativos de cuatro metodologías para el genotipado del VHC.

10 En este ensayo se comparó para 43 muestras de suero la eficiencia, fiabilidad y sensibilidad de las 3 metodologías competidoras para el subtipado del VHC, denominadas técnica comercial a (CTa; Versant HCV Genotype 2.0, LiPA, Siemens, basada en un ensayo de sonda lineal (LiPA) dirigida tanto a 5'UTR como a regiones centrales del VHC), técnica comercial b (CTb; Real Time HCV Genotype II, Abbott, basada en una PCR en tiempo real dirigida a NS5B para subtipar 1a y 1b y 5'-UTR para los otros genotipos) y el procedimiento de acuerdo con la invención que usa la tecnología conocida como GS-Junior (o plataforma de secuenciación profunda GS-Junior 454). Se tomó como referencia el procedimiento de secuenciación de Sanger aceptando su incapacidad para detectar infecciones múltiples cuando el subtipo minoritario está por debajo de un 20 %. La secuenciación de Sanger es un procedimiento de secuenciación de ADN basado en la incorporación selectiva de didesoxinucleótidos terminadores de cadena por ADN-polimerasa. Se utilizó un kit de secuenciación Bigdye terminator v1.1 cycle (Applied Biosystems) y un instrumento de secuenciación de ADN automatizada capilar (Applied Biosystems).

20 Los datos comparativos de las 43 muestras se resumen en la tabla 4.

Extracción de las muestras:

25 La sangre de los 43 individuos infectados por el VHC y los niveles detectables de ARN del VHC por encima de 100 000 copias/ml se recogió en tubos Vacutainer sin aditivos. Después de la coagulación, se obtuvo suero mediante centrifugación a baja velocidad.

Extracción de ARN vírico:

30 Se usaron 0,75 ml de suero de cada uno de los individuos sometidos a prueba para la extracción automatizada del ARN usando el kit de aislamiento total de ácidos nucleicos Total Nucleic Acid Isolation (TNAI, Roche Diagnostics) de acuerdo con las instrucciones del fabricante.

35 1. A) Generación de bibliotecas de amplicones para la reacción de secuenciación con la tecnología GS-Junior (de Roche) para llevar a cabo el procedimiento *in vitro* de la invención:

Para cada una de las 43 muestras se obtuvo un amplicón de la región 5'-UTR-Núcleo y un amplicón para la región NS5B.

40 Aquellos cebadores no disponibles comercialmente y en particular los cebadores (oligonucleótidos) de la invención se obtuvieron mediante síntesis química estándar usando cualquier kit comercial para este fin.

El rendimiento de la RT-PCR fue el siguiente:

45 A un volumen final de 50 µl se añadieron los siguientes ingredientes: 10 µl de tampón (5x) (kit Transcriptor one-step RT-PCR, Roche), 1 µl (20 µM) de cada uno de los cebadores (directo e inverso) para amplificar la 5'-UTR-Núcleo o 1,5 µl (20 µM) de cada uno de los cebadores degenerados (directo e inverso) para amplificar la región NS5B, 1 µl de polimerasa (kit Transcriptor one-step RT-PCR, Roche), 5 µl de la muestra que contiene ARN vírico (de prueba o control), los µl correspondientes de agua sin RNasa (kit Transcriptor one-step RT-PCR, Roche) hasta un volumen final de 50 µl. Las mezclas se introdujeron en el termociclador (Applied Biosystems).

50

Los ciclos de RT-PCR fueron:

Para RT 30 minutos (min) a 50 °C, y 7 min a 94 °C.

55

Para la PCR, acoplada a la fase RT, 10 ciclos de 10 segundos (s) a 94 °C, 30 s a 50 °C y 1 min a 68 °C, seguidos de 20 ciclos de 10 s a 94 °C, 30 s a 50 °C, 1 min a 68 °C (aumentando 5 segundos por ciclo), seguido de 7 min más a 68 °C para finalmente bajar la temperatura a 4 °C.

60

Las secuencias de cebadores usados específicos se indican en la tabla 1.

Tabla 1. Cebadores usados en la RT-PCR (etapa i)

Región del VHC	Cebador directo	Cebador inverso
5'-UTR-Núcleo	CTGTGAGGAAGTACTG TCTTCACGCAG (SEQ ID NO: 3)	CTAGTCGCGCGCACACCCA (SEQ ID NO: 1)
Región NS5B	Cebadores de CNTAYGAYACCMGNT GYTTTGACTC (SEQ ID NO: 7) genérica. Mezcla de oligonucleótidos que consiste en la SEQ ID NO: 54 a la SEQ ID NO: 69	Cebadores de TTNGADGAGCADGATGTW/ BAGCTC (SEQ ID NO: 5) genérica. Mezcla de oligonucleótidos que consiste en la SEQ ID NO: 36 a la SEQ ID NO: 53

5 Rendimiento de la PCR hemianidada (amplificación adicional de fragmentos de la RT-PCR para obtener una cantidad para su secuenciación) (etapa ii):

Cada una de las muestras previamente sometidas a RT-PCR se amplificó además en una PCR hemianidada con el sistema de PCR FastStart High Fidelity, dNTPack (Roche).

10 A un volumen final de 50 µl se añadieron los siguientes ingredientes: 5 µl de tampón (10x, FastStart High Fidelity de Roche), 1 µl (20 µM) de cada uno de los cebadores (directo e inverso) para amplificar la 5'-UTR-Núcleo o 1,5 µl (20 µM) de cada uno de los cebadores degenerados (directo e inverso) para amplificar la región NS5B, 0,5 µl de polimerasa (sistema de PCR FastStart High Fidelity, Roche) 1 µl (40 µM) de dNTP, 2 µl de ADN amplificado (5-250 ng), y los µl correspondientes de agua sin RNasa hasta un volumen final de 50 µl. Las mezclas se introdujeron en el termociclador (Applied Biosystems).

15 La amplificación se llevó a cabo de la siguiente manera:

20 1 min a 95 °C, seguido de 30 ciclos de 20 s a 95 °C 20 s a 55 °C, 1 min a 72 °C, y a continuación después del último ciclo 5 min a 72 °C, para finalmente bajar la temperatura a 4 °C.

25 La siguiente tabla 2 muestra los cebadores utilizados en la PCR hemianidada. Todos ellos incluyeron en su sección 5' (extremo) una secuencia de oligonucleótido universal correspondiente a un fragmento del bacteriófago M13: GTTGTA AACGACGGCCAGT (SEQ ID NO: 9) para los cebadores directos y CACAGGAAACAGCTATGACC (SEQ ID NO: 10) para los cebadores inversos.

Región del VHC	Cebador directo	Cebador inverso
5'-UTR-Núcleo	GTTGTA AACGACGGCCAGT CGGCCAGTGTCT GCGGAACCGGTGAGTACA (SEQ ID NO: 17)	CACAGGAAACAG CTATGACCCTAG TCGCGCGCACACCCA (SEQ ID NO: 18)
Región NS5B	Oligonucleótidos de secuencia genérica GTTGTA AACGACGGCCAGTCNTA YGAYACCMGNTG YTTTGACTC (SEQ ID NO: 19). Mezcla de oligonucleótidos que consiste en la SEQ ID NO: 9 unida inmediatamente en el extremo 5' de cada una de las secuencias SEQ ID NO: 54 a SEQ ID NO: 69	Oligonucleótidos de secuencia genérica CACAGGAAACAG CTATGACCGART AYCTGGTCATAG CNTCCGTGAA (SEQ ID NO: 20) Mezcla de oligonucleótidos que consiste en la SEQ ID NO: 10 unida inmediatamente en el extremo 5' de cada una de las secuencias SEQ ID NO: 70 a 73.

30 Finalmente, para identificar de forma inequívoca y fácil las muestras que se van a secuenciar usando la tecnología de secuenciación de alto rendimiento 454, las muestras amplificadas hemianidadas se sometieron a una reacción de PCR adicional con cebadores de 55 pares de bases (pb) ejemplificados en la tabla 3 y usando los reactivos del sistema de PCR FastStart High Fidelity, dNTPack (Roche):

Tabla 3.

Identificación	Cebadores universales con adaptador <i>oligo A o B</i> + <i>clave</i> + <i>MID</i> + <i>secuencia de oligonucleótidos del bacteriófago M13 para cebadores directos (Fw) o inversos (Rv)</i>
AFw1M13u (SEQ ID NO: 21)	<u>CGTATCGCCTCCCTCGCGCCATCAGACGAGTGCGT</u> <u>GTTGTA</u> <u>AAAACGACGGCCAGT</u>
BRv1M13d (SEQ ID NO: 22)	<u>CTATGCGCCTTGCCAGCCCGCTCAGACGAGTGCGT</u> <u>CACAGGAAACAGCTATGACC</u>
AFw2M13u (SEQ ID NO: 23)	<u>CGTATCGCCTCCCTCGCGCCATCAGACGCTCGACAGTTGTA</u> <u>AAAACGACGGCCAGT</u>
BRv2M13d (SEQ ID NO: 24)	<u>CTATGCGCCTTGCCAGCCCGCTCAGACGCTCGACACACAGGAAACAGCTATGACC</u>
AFw3M13u (SEQ ID NO: 25)	<u>CGTATCGCCTCCCTCGCGCCATCAGAGACGCACTCGTTGTA</u> <u>AAAACGACGGCCAGT</u>
BRv3M13d (SEQ ID NO: 26)	<u>CTATGCGCCTTGCCAGCCCGCTCAGAGACGCACTCCACAGGAAACAGCTATGACC</u>
AFw4M13u (SEQ ID NO: 27)	<u>CGTATCGCCTCCCTCGCGCCATCAGAGCACTGTAGGTTGTA</u> <u>AAAACGACGGCCAGT</u>
BRv4M13d (SEQ ID NO: 28)	<u>CTATGCGCCTTGCCAGCCCGCTCAGAGCACTGTAGCACAGGAAACAGCTATGACC</u>
AFw5M13u (SEQ ID NO: 29)	<u>CGTATCGCCTCCCTCGCGCCATCAGATCAGACACGGTTGTA</u> <u>AAAACGACGGCCAGT</u>
BRv5M13d (SEQ ID NO: 30)	<u>CTATGCGCCTTGCCAGCCCGCTCAGATCAGACACGCACAGGAAACAGCTATGACC</u>

- 5 CGTATCGCCTCCCTCGCGCCA (SEQ ID NO: 13) corresponde al adaptador 5' conocido como oligo A que se adherirá al oligonucleótido complementario en las microesferas para la secuenciación.
- CTATGCGCCTTGCCAGCCCGC (SEQ ID NO: 14) corresponde al adaptador 5' conocido como oligo B que se adherirá al oligonucleótido complementario en las microesferas para la secuenciación.
- 10 TCAG corresponde a una secuencia clave de 4 nucleótidos que se identifica mediante el secuenciador del genoma (GS-Junior o GS-FLX, Roche) durante el procedimiento de secuenciación. Esta clave es esencial para validar y normalizar cada secuencia durante el experimento con GS-Junior.
- 15 GTTGTA AAAACGACGGCCAGT (SEQ ID NO: 9) corresponde a la secuencia de fragmentos del bacteriófago M13 para los cebadores directos.
- CACAGGAAACAGCTATGACC (SEQ ID NO: 10) corresponde a la secuencia de fragmentos del bacteriófago M13 para cebadores inversos.
- 20 Las secuencias ACGAGTGCGT (SEQ ID NO: 31), ACGCTCGACA (SEQ ID NO: 32), AGACGCACTC (SEQ ID NO: 33), AGCACTGTAG (SEQ ID NO: 34) y ATCAGACACG (SEQ ID NO: 35) también conocidas como MID, que significa "Identificadores moleculares", son secuencias cortas, validadas por el fabricante (Roche), que se usan como un identificador del paciente. Las amplificaciones por RT-PCR anidada de cada paciente se mezclan en la biblioteca monocatenaria antes de iniciar el experimento de secuenciación GS-Junior. Al finalizar el experimento,
- 25 se informa de un archivo con una mezcla compleja de secuencias. El MID permite identificar y separar las secuencias de cada paciente. Los cebadores mostrados en la tabla 3 corresponden a un tipo representativo (5 MID) de pares de cebadores utilizados para 5 de las muestras analizadas. De hecho, se usaron 43 MID diferentes para las 43 muestras diferentes sometidas a prueba. Estas secuencias cortas las proporcionaron los proveedores (Roche) que aseguran que los 43 MID identifican inequívocamente, sin riesgo de confusión, las muestras deseadas que van a someterse a prueba.
- 30 La mezcla de reacción incluyó hasta un volumen final de 50 µl los siguientes ingredientes: 5 µl de tampón (10x), 1 µl (20 µM) de cada uno de los cebadores (directo e inverso) para amplificar la 5'-UTR-Núcleo o 1,5 µl (20 µM) de cada uno de los cebadores degenerados directo e inverso, 1 µl (40 µM) de dNTP, 2 µl de ADN amplificado (5-250 ng) y los µl correspondientes de agua sin RNasa (kit Transcriptor one-step RT-PCR, Roche) hasta un volumen final de
- 35 50 µl. Las mezclas se introdujeron en el termociclador (Applied Biosystems).
- Una vez amplificados, los fragmentos que se van a secuenciar adicionalmente y correspondientes a la región 5'-UTR-Núcleo abarcada tienen, en general, una longitud mínima de 446 nucleótidos debido a las diferentes secuencias acopladas. De la misma manera, los fragmentos que se van a secuenciar adicionalmente y que
- 40 corresponden a la región NS5B abarcada tienen, en general, una longitud mínima de 454 nucleótidos debido a las diferentes secuencias acopladas.

5 Todos los productos de la RT-PCR anidada, una vez purificados, se mezclaron (combinaron) y se secuenciaron siguiendo el protocolo de la tecnología de secuenciación de alto rendimiento 454, obteniendo miles de secuencias (entre 60 000 y 100 000) para cada muestra. Se usó un algoritmo interno (expuesto brevemente a continuación) para separar las lecturas secuenciadas para obtener un archivo fasta separado con las lecturas correspondientes a cada
 10 muestra y amplicón (paciente y región del VHC, ya sea 5'-UTR-Núcleo o NS5B). Esta separación se lleva a cabo emparejando los correspondientes MID y cebadores por alineación con cada lectura secuenciada. Se obtuvieron cientos de secuencias para cada paciente y para cada amplicón. Las secuencias se compararon filogenéticamente con las referencias de secuencia obtenidas de GeneBank para la clasificación de los subtipos. Este algoritmo interno puede detectar si el paciente está infectado por más de un subtipo al mismo tiempo. Los resultados de identificación o diagnóstico usando el procedimiento *in vitro* de la invención se enumeran en la tabla 4 en la columna identificada como GS-Junior.

15 Brevemente, en el algoritmo, las secuencias en un archivo fasta obtenidas del GS-454 Junior se separaron identificando MID y cebador, generando un solo archivo fasta para cada muestra, región y cadena. Seguía un filtro de calidad en el que se descartaron las secuencias que tenían más de un número dado de interrupciones, o más de un número dado de Ns (nucleótido no definido A, C, G, T), o más de un número dado de diferencias respecto del haplotipo dominante (secuencia predominante en la muestra). Para las secuencias aceptadas, las interrupciones y Ns restantes se corrigieron según el contenido del haplotipo dominante. Las cadenas complementarias inversas
 20 (RV) se mezclaron con las cadenas directas (FW) para obtener un único conjunto de haplotipos y frecuencias para cada muestra y región. Se obtuvo una estimación de la variabilidad de cuasiespecies a partir del conjunto completo de haplotipos con las frecuencias correspondientes por medio de la entropía de Shannon (Sn) y la diversidad nucleotídica (Pi). A continuación se seleccionaron los haplotipos que se utilizarán para el subtipado. Estos fueron el haplotipo dominante y todos aquellos haplotipos con una identidad respecto del haplotipo dominante por debajo de un umbral dado, que es dependiente de la región. Los haplotipos seleccionados se alinearon de forma múltiple con las referencias de subtipos de la región correspondiente (5'-UTR-Núcleo o NS5B). A partir de la alineación múltiple y el modelo de mutación óptima de la región, se obtuvo una matriz de distancias genéticas entre las referencias y los haplotipos seleccionados para el paciente. Estas distancias se usaron luego para construir un árbol filogenético UPGMA y para estimar el subtipo más cercano a cada uno de los haplotipos seleccionados del paciente. También se
 30 obtuvo un nivel de confianza por medio de un muestreo con reposición (*bootstrap*) en la alineación múltiple.

Con fines ilustrativos, se muestra un modo de salida o presentación de los resultados del algoritmo en las figuras 1 y 2.

35 Estas figuras 1 y 2 muestran, respectivamente, el diagnóstico del paciente de la muestra 43 (P43). Esta muestra se diagnosticó como coinfectada por los subtipos 4d y 1a del VHC. Un diagnóstico gráfico de la infección por el subtipo 1a se ilustra en la figura 1 (y su continuación Cont.), en el que la secuencia amplificada del VHC en la muestra P43 (Px10.66rd) está filogenéticamente próxima al subtipo 1a (referencia 1a LTD1AF51). Esta correlación se realiza al determinar la distancia entre subtipos dispuestos en ramas en un árbol filogenético. Además, la muestra 43 también dio lugar a la amplificación significativa de otra secuencia (Px1, 1506rd), que se correlacionaba bien con el subtipo 4d (referencia 4dDQ418786) como se puede ver gráficamente en la figura 2 (y su continuación Cont.). Este tipo de salidas o representaciones permiten determinar fácilmente la presencia de múltiples infecciones por el VHC en una muestra. Cada una de la figura 1 y la figura 2 en su conjunto muestra todo el diagrama del árbol filogenético de los diferentes subtipos del VHC. A efectos de visualización, ambas figuras (figura 1 y figura 2) se han cortado en dos partes (figuras parciales) y están dispuestas para que la figura completa se pueda ensamblar sin ocultar ninguna parte de las figuras parciales.

1. B) Diagnóstico de las 43 muestras utilizando la técnica comercial a (CTa; Versant HCV Genotype 2.0, LiPA, Siemens).

50 Esta técnica se basa en un ensayo de sonda lineal (LiPA) dirigido a las regiones tanto 5'-UTR como central del VHC, y se realizó siguiendo las instrucciones del producto comercial. El resultado notificado por esta técnica para clasificar al paciente 43 es un resultado indeterminado que necesitará una mayor validación.

1. C) Diagnóstico de las 43 muestras utilizando la técnica comercial b (CTb; Real Time HCV Genotype II, Abbott).

Esta técnica se basa en una PCR en tiempo real dirigida a NS5B para subtipar 1a y 1b y 5'UTR para los otros genotipos. Se realizó siguiendo las instrucciones del kit comercial. El resultado notificado por esta técnica para clasificar al paciente 43 es un resultado indeterminado que necesitará una mayor validación.

60 1. D) La secuenciación de Sanger es un procedimiento de secuenciación de ADN basado en la incorporación selectiva de didesoxinucleótidos terminadores de cadena por ADN-polimerasa. Utilizamos el kit de secuenciación Bigdye terminator v1.1 cycle (Applied Biosystems) y un instrumento de secuenciación de ADN automatizado capilar (Applied Biosystems). El problema de esta técnica que se ha considerado como la referencia para la clasificación del VHC es que no puede detectar si un paciente está infectado por más de un subtipo al mismo tiempo y clasifica al paciente 43 como 4d, que es el subtipo más representado en el paciente.

ES 2 676 245 T3

Como se indicó anteriormente, la siguiente tabla 4 muestra los resultados del diagnóstico con los cuatro procedimientos analizados en los que se usa el procedimiento de Sanger como referencia.

5 Tabla 4. Resultados comparativos usando los cuatro procedimientos.

Orden	Sanger	CTa	CTb	GS-Junior (invención)
1	1a	1a	1a	1a
2	1a	1a	1a	1a
3	1a	1a	1a	1a
4	1b	1b	1	1b
5	1b	1b	1b	1b
6	1b	1b	1	1b
7	1b	1b	1b	1b
8	1b	1b	1b	1b
9	1b	1b	1b	1b
10	1b	1b	1b	1b
11	2c	2a/2c	2	2c
12	2c	2a/2c	2	2c
13	3a	3	3	3a
14	3a	3a	ND	3a
15	3a	3a	3	3a
16	3a	3a	3	3a
17	4a	4	4	4a
18	4a	4a/c/d	4	4a
19	4d	4	4	4d
20	4d	4	4	4d
21	4d	4a/c/d	4	4d
22	4d	4a/c/d	4	4d
28	4f	4f	4 + 5	4f
29	4f	4f	4 + 5	4f
23	4f	1 b + 4f (4 no claro)	4 + 5	4f
24	4f	1 b + 4f (4 no claro)	4 + 5	4f
25	4f	1 b + 4f (4 no claro)	4 + 5	4f
26	4f	1 b + 4f (4 no claro)	4 + 5	4f
27	4f	1 b + 4f (4 no claro)	4 + 5	4f
33	4f	Ind (5, 9, 20)	4 + 5	4f
35	4f	Ind (5,9,16)	4 + 5	4f
30	4f	Ind (5, 9, 20)	5	4f
31	4f	Ind (5, 9, 20)	5	4f
32	4f	Ind (5, 9, 20)	5	4f
34	4f	Ind (5, 9, 20)	5	4f
36	4r	1 + 4e (no claro)	1 + 4	4r
37	5a	5a	5	5a
38	6c	neg	5	6c
39	2j	2k	ND	2j+4f
40	4p	3a + 4a/c/d	3 + 4	3a+4p
41	4d	4	4	4d+1b
42	1a	1+4	4	4d+1a
43	4d	1 a + 4a/c/d	1a + 4	4d+1a

ND significa no determinado; Ind significa indeterminado; neg. significa negatividad para la amplificación.

5 Como se puede ver en esta tabla 4, la CTa pudo clasificar los subtipos 1a, 1b, y 3a y 5a, pero no pudo subtipar el resto de los genotipos (2, 4 o 6). La CTa detectó los subtipos 4a, 4c y 4d pero el informe dio una clasificación indeterminada, que se notificó como 4 en algunos casos, pero como 4a/c/d en otros. La CTa no pudo subtipar la mayoría del 4f (excepto el caso 28, 29) o 6, lo que dio un resultado de negatividad para el genotipo 6 en la muestra 38.

10 Por otro lado, los resultados utilizando la técnica comercial b (CTb) mostraron que esta técnica no podía subtipar los genotipos 2, 3, 4, 5 o 6. La CTb pudo diferenciar entre subtipo 1a y 1b, pero falló en dos (muestras 4 y 6) de 7 muestras de subtipo 1b (muestras 4-10) que la CTb clasificó como 1 no definido y en un caso no pudo diferenciar entre coinfecciones por 1b, 1a y 4 (muestras 41, 42 y 43). La metodología de la CTb no pudo identificar ningún subtipo 4f. Cuatro de 13 subtipos 4f se clasificaron por la CTb como genotipo 5 (muestras 30, 31, 32 y 34) y el resto se clasificó como 4 o 5 (4+5) (muestras 23, 24, 25, 26, 27, 28, 29, 33 y 35). La CTb clasificó el subtipo 4r (muestra 15 36) como 1 o 4. El subtipo 6 fue erróneamente clasificado como 5.

20 Finalmente, los resultados con el procedimiento *in vitro* de acuerdo con la invención (usando tecnología de la plataforma de secuenciación profunda GS-Junior 454) revelaron que el procedimiento *in vitro* de la invención pudo subtipar las 43 muestras. GS-Junior (para aplicar el procedimiento *in vitro* de la invención) detectó 5 muestras infectadas de forma múltiple (coinfecciones) por más de un subtipo, identificando la proporción de cada subtipo (muestras 39, 40, 41, 42 y 43). Por lo tanto, el procedimiento *in vitro* de la invención incluso permitió la detección de una infección múltiple no detectable habitualmente por el procedimiento de referencia (Sanger), debido al hecho de que el procedimiento de referencia da para cada muestra analizada un promedio de las secuencias detectadas.

25 El resultado final de la plataforma de secuenciación profunda GS-Junior 454 fue un archivo Fasta con cientos o miles de secuencias que pueden utilizarse para nuevos estudios o reclasificaciones futuras si es necesario como se detalla anteriormente.

30 Todos estos datos considerados conjuntamente permiten concluir que el procedimiento *in vitro* de la invención es fácil de realizar y, por lo tanto, comercialmente aplicable en forma de kits. Además, el procedimiento de la invención permite la detección correcta de muestras coinfectadas, así como la detección de infecciones por virus recombinantes.

35 Es deducible directamente de los resultados que el procedimiento de la invención es más fiable y sensible que otras técnicas comerciales en términos de que permite una identificación inequívoca del genotipo y subtipo del VHC en algunas muestras de complejos que no se pueden resolver mediante otras técnicas. El procedimiento de la invención además es más preciso que las técnicas de la técnica anterior, ya que no solo permite la resolución de múltiples muestras infectadas, sino también la identificación de los subtipos entre los múltiples genotipos identificados en estas muestras infectadas de forma múltiple.

40 Como se indica a lo largo de esta descripción, determinar las variantes específicas en pacientes infectados por el VHC es de gran interés para diagnosticar al paciente de una manera más precisa, ya que de un diagnóstico exacto se deriva un tratamiento seleccionado eficaz.

45 Para ilustrar el significado de un diagnóstico exacto, el próximo ejemplo 2 muestra el efecto del subtipo del VHC detectado sobre un tratamiento específico.

Ejemplo 2. Datos indicativos del efecto del subtipo del VHC sobre el resultado del tratamiento con interferón pegilado y rivabirina (pegIFN+RBV)

50 La eficacia de la respuesta virológica mantenida (RVM), que se define como el ARN del VHC en suero indetectable seis meses después de la finalización del tratamiento, está directamente relacionada con el genotipo. Las tasas de RVM van de un 38 a un 41 % para el genotipo 1 (Hadziyannis SJ et al. for PEGASYS International Study Group, "Peginterferonalpha 2a and ribavirin combination therapy in chronic hepatitis C: a randomized study of treatment duration and ribavirin dose", Ann Intern Med - 2004. vol. 140, pp.: 346-355), un 93 % para el genotipo 2, un 79 % para el genotipo 3 (Zeuzem S, et al., "Peginterferon alfa-2b plus ribavirin for treatment of chronic hepatitis C in previously untreated patients infected with HCV genotypes 2 or 3", J Hepatol - 2004, vol. 40, pp.: 993-9) y un 69 % para el genotipo 4 (Kamal SM et al., "Peginterferon -2b and ribavirin therapy in chronic hepatitis C genotype 4: impact of treatment duration and viral kinetics on sustained virological response. Gut - 2005, vol. 54, pp.: 858-66).

60 Sin embargo, la falta de un sistema altamente preciso para clasificar a los pacientes según el subtipo del VHC ha impedido evaluar la respuesta a pegIFN + RBV asociada a cada subtipo particular.

65 La tabla 5 siguiente resume los resultados de un estudio preliminar de la asociación del subtipo con el resultado final después del tratamiento con el tratamiento de referencia con pegIFN+RBV

Existen algunos datos indicativos de que los pacientes infectados por el subtipo 2b y 4d tuvieron un grado particularmente bajo de respuesta a pegIFN+RBV. Estos datos preliminares sugieren que el subtipado puede ser un valor pronóstico importante antes de iniciar cualquier tratamiento en pacientes con infección crónica por el VHC.

5 Tabla 5. Estudio de la asociación del subtipo y la respuesta a pegIFN+RBV

		RESPUESTA A pegIFN+RBV			
		RVM		NNR (Nulo, Rebrote o Recaída)	
SUBTIPO	TOTAL número de pacientes	Número de pacientes	Porcentaje	Número de pacientes	Porcentaje
1a	6	2	33,33 %	4	66,67 %
1b	17	8	47,06 %	9	52,94 %
2b	2	0	0,00 %	2	100,00 %
2c	5	3	60,00 %	2	40,00 %
3a	8	6	75,00 %	2	25,00 %
4a	2	1	50,00 %	1	50,00 %
4d	6	1	16,67 %	5	83,33 %

10 NR significa Ninguna respuesta. Este tipo de NR puede ser de diferentes tipos: a) Nulo, cuando desde el inicio del tratamiento la carga vírica no baja; Rebrote virológico (Rebrote) cuando en las etapas iniciales del tratamiento la carga vírica disminuye pero aparece después; y Recaída, cuando la carga vírica se elimina (o se reduce a cantidades muy pequeñas e indetectables) dentro del tratamiento, pero después de dejar dicho tratamiento, la carga vírica aparece nuevamente aproximadamente 6 meses después al interrumpir el tratamiento.

15 Por lo tanto, con el objetivo de resolver los costes y el tiempo (realmente importante para los pacientes), está muy previsto identificar con exactitud el genotipo y subtipo de la variante del VHC en una muestra aislada infectada. El procedimiento de la invención, que usa los oligonucleótidos específicos para abarcar las regiones variantes significativas y distintivas del VHC clave, se concibe como una herramienta útil para llevar a cabo este objetivo.

REFERENCIAS CITADAS EN LA SOLICITUD

- 5 - Simmonds et al., "Consensus proposals for a unified system of nomenclature of hepatitis C virus genotypes", Hepatology - 2005, vol. 42, pp. 963-973.
- 10 - Chevaliez S. et. al. "Hepatitis C virus (HCV) genotype 1 subtype identification in new HCV drug development and future clinical practice", PLoS ONE, 2009, vol. 4, pp. 1-9.
- 10 - Kuiken et al., "A Comprehensive System for Consistent Numbering of HCV Sequences, Proteins and Epitopes", Hepatology - 2006, DOI 10.1002/hep.21377, pp. 1355-1361.
- 15 - Lauck M., et. al. "Analysis of hepatitis C virus intra-host diversity across the coding region by ultra-deep sequencing". J. Virol. - 2012, Epub (publicación en línea, antes de la impresión).
- 15 - Hadziyannis SJ et al. for PEGASYS International Study Group, "Peginterferonalpha 2a and ribavirin combination therapy in chronic hepatitis C: a randomized study of treatment duration and ribavirin dose", Ann Intern Med - 2004, vol. 140, pp.: 346-355.
- 20 - Zeuzem S, et al., "Peginterferon alfa-2b plus ribavirin for treatment of chronic hepatitis C in previously untreated patients infected with HCV genotypes 2 or 3", J Hepatol - 2004, vol. 40, pp.: 993-9.
- 20 - Kamal SM et al., "Peginterferon-2b and ribavirin therapy in chronic hepatitis C genotype 4: impact of treatment duration and viral kinetics on sustained virological response. Gut - 2005, vol. 54, pp.: 858-66.

ES 2 676 245 T3

LISTADO DE SECUENCIAS

5	<110>	Fundació Hospital Universitari Vall d'Hebron - Institut de Recerca CIBER (CENTRO DE INVESTIGACIÓN BIOMÉDICA EN RED) ROCHE DIAGNOSTICS, S.L.	
	<120>	Cebadores y procedimientos para detectar variantes del virus de la hepatitis C (VHC) humano en una muestra aislada	
10	<130>	P2508PC00	
	<150>	EP13382278	
	<151>	05/07/2013	
15	<160>	73	
	<170>	PatentIn versión 3.5	
20	<210>	1	
	<211>	19	
	<212>	ADN	
	<213>	Artificial	
25	<220>		
	<223>	Cebador inverso	
	<400>	1	
		ctagtcgctg gcacaccca	19
30	<210>	2	
	<211>	914	
	<212>	ADN	
	<213>	Virus de la hepatitis C	
35	<400>	2	
		gccagccccc tgatgggggc gacactccac catgaatcac tcccctgtga ggaactactg	60
		tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgcctgacag cctccaggac	120
		ccccctccc gggagagcca tagtggctcg cggaaccggt gactacaccg gaattgccag	180
		gacgaccggg tcctttcttg gataaaccgc ctcaatgcct ggagatttgg gcgtgcccc	240
		gcaagactgc tagccgagta gtgtgggtc gcgaaaggcc ttgtgtact gcctgatagg	300
		gtgcttgcca gtgccccggg aggtctcgta gaccgtgac catgagcac aatcctaac	360
		ctcaaagaaa aaccaaactg aacaccaacc gtcgcccaca ggacgtcaag ttcccgggtg	420
		gcggtcagat cgttgggtga gttacttgt tgccgcgac gggccctaga ttgggtgtgc	480
		gcgcgacgag gaagacttcc gagcggctgc aacctcgagg tagacgtcag cctatcccca	540
		aggcacgtcg gcccgagggc aggacctggg ctacgcccgg gtacccttgg cccctctatg	600
		gcaatgaggg ttgccccggg gcgggatggc tcctgtctcc ccgtggctct cggcctagct	660
		ggggccccac agacccccgg cgtaggtcgc gcaatttggg taaggtcatc gataccctta	720
		cgtgcccgtt cgcggacctc atggggtaca taccgctcgt cggcgcccct cttggaggcg	780

ES 2 676 245 T3

ctgccagggc cctggcgcat ggcgtccggg ttctggaaga cggcgtgaac tatgcaacag 840
 ggaaccttcc tggttgctct ttctctatct tccttctggc cctgctctct tgcctgactg 900
 tgcccgcttc agcc 914
 <210> 3
 <211> 27
 5 <212> ADN
 <213> Artificial
 <220>
 <223> Cebador directo
 10 <400> 3
 ctgtgaggaa ctactgtctt cacgcag 27
 <210> 4
 <211> 22
 <212> ADN
 <213> Artificial
 <220>
 <223> Cebador directo
 20 <400> 4
 gtctgaggaa ccggtgagta ca 22
 <210> 5
 <211> 26
 <212> ADN
 <213> Artificial
 <220>
 <223> Cebador inverso
 30 <220>
 <221> misc_feature
 <222> (3)..(3)
 <223> n es a, c, g, o t
 35 <220>
 <221> misc_feature
 <222> (6)..(6)
 <223> D signfica G, A o T
 40 <220>
 <221> misc_feature
 <222> (12)..(12)
 <223> D signfica G, A o T
 45 <220>
 <221> misc_feature
 <222> (18)..(18)
 <223> W signfica A o T
 50 <220>
 <221> misc_feature
 <222> (21)..(21)
 <223> B signfica G, T o C
 55 <400> 5
 ttngadgagc adgatgtwat bagctc 26
 60 <210> 6

ES 2 676 245 T3

<211> 1776
 <212> ADN
 <213> Virus de la hepatitis C

5 <400> 6
 tcaatgtctt attcctggac aggcgcactc gtcaccccgt gcgctgcgga agaacaaaaa 60
 ctgccatca acgcactgag caactcgttg ctacgccatc acaatctggt gtattccacc 120
 acttcacgca gtgcttgcca aaggcagaag aaagtacatc ttgacagact gcaagttctg 180
 gacagccatt accaggacgt gctcaaggag gtcaaagcag cggcgtcaaa agtgaaggct 240
 aacttgctat ccgtagagga agcttgcacg ctgacgcccc cacattcagc caaatccaag 300
 tttggctatg gggcaaaaga cgtccgttgc catgccagaa aggccgtagc ccacatcaac 360
 tccgtgtgga aagaccttct ggaagacagt gtaacaccaa tagacactac catcatggcc 420
 aagaacgagg ttttctgctg tcagcctgag aaggggggtc gtaagccagc tcgtctcatc 480
 gtgttccccg acctgggctg gcgctgtgtc gagaagatgg ccctgtacga cgtggtttagc 540
 aagctcccc tggccgtgat ggggaagctcc tacggattcc aatactcacc aggacagcgg 600
 gttgaattcc tcgtgcaagc gtggaagtcc aagaagacc c gatgggggtt ctogtatgat 660
 acccgctgtt ttgactccac agtactgag agcgacatcc gtacggagga ggcaatttac 720
 caatgttggt acctggaccc ccaagcccgc gtggccatca agtccctcac tgagaggctt 780
 tatgttgggg gccctcttac caattcaagg ggggaaaact gcggctaccg cagggtgccgc 840
 gcgagcggcg tactgacaac tagctgtggt aacaccctca cttgctacat caaggcccgg 900
 gcagcctgtc gagccgcagg gctccaggac tgcaccatgc tcgtgtgtgg cgacgactta 960
 gtcgttatct gtgaaagtgc ggggggtccag gaggacgcgg cgagcctgag agccttcacg 1020
 gaggctatga ccaggtactc cgccccccc ggggacccc cacaaccaga atacgacttg 1080
 gagctataa catcatgctc ctccaacgtg tcagtcgccc acgacggcgc tggaaagagg 1140
 gtctactacc ttaccctgta ccctacaacc cccctcgca gagccgcgtg ggagacagca 1200
 agacacactc cagtcaattc ctggctagc aacataatca tgtttgcccc cacactgtgg 1260
 gcgaggatga tactgatgac ccatttcttt agcgtcctca tagccaggga tcagctttaa 1320
 caggctctta actgtgagat ctacggagcc tgctactcca tagaaccact ggatctacct 1380
 ccaatcattc aaagactcca tggcctcagc gcattttcac tccacagtta ctctccaggt 1440
 gaaatcaata ggggtggccgc atgcctcaga aaacttgggg tcccgcctt gcgagcttgg 1500
 agacaccggg cccggagcgt ccgcgctagg cttctgtcca gaggaggcag ggctgccata 1560

ES 2 676 245 T3

	tgtggcaagt acctcttcaa ctgggcagta agaacaaagc tcaaactcac tccaatagcg	1620
	gccgctggcc ggctggactt gtccggttgg ttcacggctg gctacagcgg gggagacatt	1680
	tatcacagcg tgtctcatgc ccggccccgc tggttctggt tttgcctact cctgctcgct	1740
	gcaggggtag gcattctacct cctccccaac cgatga	1776
5	<210> 7 <211> 25 <212> ADN <213> Artificial	
10	<220> <223> Cebador directo	
15	<220> <221> misc_feature <222> (2)..(2) <223> n es a, c, g, o t	
20	<220> <221> misc_feature <222> (5)..(5) <223> Y significa T o C	
25	<220> <221> misc_feature <222> (8)..(8) <223> Y significa T o C	
30	<220> <221> misc_feature <222> (12)..(12) <223> M significa A o C	
35	<220> <221> misc_feature <222> (14)..(14) <223> n es a, c, g, o t	
40	<220> <221> misc_feature <222> (17)..(17) <223> Y significa T o C	
	<400> 7	
	cntaygayac cmgntgyttt gactc	25
45	<210> 8 <211> 26 <212> ADN <213> Artificial	
50	<220> <223> Cebador inverso	
55	<220> <221> misc_feature <222> (3)..(3) <223> R significa G o A	
	<220> <221> misc_feature <222> (6)..(6)	

ES 2 676 245 T3

<223> Y significa T o C
 <220>
 <221> misc_feature
 5 <222> (18).. (18)
 <223> n es a, c, g, o t
 <400> 8
gartayctgg tcatagcntc cgtgaa 26
 10
 <210> 9
 <211> 20
 <212> ADN
 <213> Bacteriófago M13
 15 <400> 9
gttgtaaaac gacggccagt 20
 <210> 10
 20 <211> 20
 <212> ADN
 <213> Bacteriófago M13
 <400> 10
 25 **cacaggaaac agctatgacc** 20
 <210> 11
 <211> 41
 30 <212> ADN
 <213> Artificial
 <220>
 <223> Fragmento de un cebador directo que incluye un fragmento del bacteriófago M13 + adaptador universal
 35 oligo A para la tecnología de pirosecuenciación 454
 <400> 11
cgtatcgcct ccctcgcgcc agttgtaaaa cgacggccag t 41
 <210> 12
 40 <211> 41
 <212> ADN
 <213> Artificial
 <220>
 45 <223> Fragmento de un cebador inverso que incluye un fragmento del bacteriófago M3 + adaptador universal
 oligo B para la tecnología de pirosecuenciación 454
 <400> 12
 50 **ctatcgcct tgccagcccg ccacaggaaa cagctatgac c** 41
 <210> 13
 <211> 21
 <212> ADN
 <213> Artificial
 55 <220>
 <223> Adaptador oligo A para cebadores directos en la tecnología de pirosecuenciación 454
 <400> 13
 60 **cgtatcgcct ccctcgcgcc a** 21
 <210> 14
 <211> 21

ES 2 676 245 T3

	<212>	ADN	
	<213>	Artificial	
	<220>		
5	<223>	Adaptador oligo B para cebadores inversos en la tecnología de pirosecuenciación 454	
	<400>	14	
		ctatgcgct tgccagcccg c	21
10	<210>	15	
	<211>	20	
	<212>	ADN	
	<213>	Artificial	
15	<220>		
	<223>	Cebador inverso	
	<220>		
	<221>	misc_feature	
20	<222>	(1)..(1)	
	<223>	N significa (N)m, en el que N es cualquier nucleótido seleccionado de A, T, C o G, y m es un número entero de 0 a 25	
	<400>	15	
25		nctagtgcg cgcacaccca	20
	<210>	16	
	<211>	27	
	<212>	ADN	
30	<213>	Artificial	
	<220>		
	<223>	Cebador inverso	
35	<220>		
	<221>	misc_feature	
	<222>	(1)..(1)	
	<223>	N significa (N)m, en el que N es cualquier nucleótido seleccionado de A, T, C o G, y m es un número entero de 0 a 25	
40	<220>		
	<221>	misc_feature	
	<222>	(4)..(4)	
	<223>	n es a, c, g, o t	
45	<220>		
	<221>	misc_feature	
	<222>	(7)..(7)	
	<223>	D significa G, A o T	
50	<220>		
	<221>	misc_feature	
	<222>	(13)..(13)	
	<223>	D significa G, A o T	
55	<220>		
	<221>	misc_feature	
	<222>	(19)..(19)	
	<223>	W significa A o T	
60	<220>		
	<221>	misc_feature	
	<222>	(22)..(22)	
	<223>	B significa G, T o C	
65	<220>		

ES 2 676 245 T3

	<400> 16		
	nttngadgag cadgatgtwa tbagctc		27
5	<210> 17 <211> 42 <212> ADN <213> Artificial		
10	<220> <223> Cebador directo		
	<400> 17		
	gttgtaaaac gacggccagt gtctgaggaa ccggtgagta ca		42
15	<210> 18 <211> 39 <212> ADN <213> Artificial		
20	<220> <223> Cebador inverso		
	<400> 18		
	cacaggaaac agctatgacc ctagtgcgac gcacaccca		39
25	<210> 19 <211> 45 <212> ADN <213> Artificial		
30	<220> <223> Cebador directo		
35	<220> <221> misc_feature <222> (22)..(22) <223> n es a, c, g, o t		
40	<220> <221> misc_feature <222> (25)..(25) <223> Y significa T o C		
45	<220> <221> misc_feature <222> (28)..(28) <223> Y significa T o C		
50	<220> <221> misc_feature <222> (32)..(32) <223> M significa A o C		
55	<220> <221> misc_feature <222> (34)..(34) <223> n es a, c, g, o t		
60	<220> <221> misc_feature <222> (37)..(37) <223> Y significa T o C		
65	<400> 19		
	gttgtaaaac gacggccagt cntaygayac cmgntgyttt gactc		45

ES 2 676 245 T3

<210> 20
 <211> 46
 <212> ADN
 5 <213> Artificial

 <220>
 <223> Cebador inverso

 10 <220>
 <221> misc_feature
 <222> (23)..(23)
 <223> R significa A o G

 15 <220>
 <221> misc_feature
 <222> (26)..(26)
 <223> Y significa T o C

 20 <220>
 <221> misc_feature
 <222> (38)..(38)
 <223> N significa A, C, G o T

 25 <400> 20
 cacaggaaac agctatgacc gartayctgg tcatagcntc cgtgaa 46

 <210> 21
 <211> 55
 30 <212> ADN
 <213> Artificial

 <220>
 <223> Cebador directo

 35 <400> 21
 cgtatcgctt ccctcgcgcc atcagacgag tgcgtgttgt aaaacgacgg ccagt 55

 <210> 22
 <211> 55
 40 <212> ADN
 <213> Artificial

 <220>
 45 <223> Cebador inverso

 <400> 22
 ctatgcgctt tgccagcccg ctcagacgag tgcgtcacag gaaacagcta tgacc 55

 50 <210> 23
 <211> 55
 <212> ADN
 <213> Artificial

 55 <220>
 <223> Cebador directo

 <400> 23
 cgtatcgctt ccctcgcgcc atcagacgct cgacagttgt aaaacgacgg ccagt 55

 60 <210> 24
 <211> 55
 <212> ADN
 <213> Artificial

ES 2 676 245 T3

<220>
 <223> Cebador inverso
 5 <400> 24
 ctatg~~gc~~cct tgccagccc~~g~~ ctcagagcgt cgacacacag gaaacagcta tgacc 55
 <210> 25
 <211> 55
 10 <212> ADN
 <213> Artificial
 <220>
 <223> Cebador directo
 15 <400> 25
 cgtatgc~~gc~~cct ccctcgcgcc atcagagacg cactcgttgt aaaacgacgg ccagt 55
 <210> 26
 20 <211> 55
 <212> ADN
 <213> Artificial
 <220>
 25 <223> Cebador inverso
 <400> 26
 ctatg~~gc~~cct tgccagccc~~g~~ ctcagagacg cactccacag gaaacagcta tgacc 55
 30 <210> 27
 <211> 55
 <212> ADN
 <213> Artificial
 35 <220>
 <223> Cebador directo
 <400> 27
 cgtatgc~~gc~~cct ccctcgcgcc atcagagcac tgtaggttgt aaaacgacgg ccagt 55
 40 <210> 28
 <211> 55
 <212> ADN
 <213> Artificial
 45 <220>
 <223> Cebador inverso
 <400> 28
 50 ctatg~~gc~~cct tgccagccc~~g~~ ctcagagcac tgtagcacag gaaacagcta tgacc 55
 <210> 29
 <211> 55
 <212> ADN
 55 <213> Artificial
 <220>
 <223> Cebador directo
 60 <400> 29
 cgtatgc~~gc~~cct ccctcgcgcc atcagatcag acacggttgt aaaacgacgg ccagt 55
 <210> 30
 <211> 55

ES 2 676 245 T3

	<212>	ADN	
	<213>	Artificial	
	<220>		
5	<223>	Cebador inverso	
	<400>	30	
		ctatgcgct tgccagcccg ctcagatcag acacgcacag gaaacagcta tgacc	55
10	<210>	31	
	<211>	10	
	<212>	ADN	
	<213>	Artificial	
15	<220>		
	<223>	Identificador molecular	
	<400>	31	
		acgagtgcgt	10
20	<210>	32	
	<211>	10	
	<212>	ADN	
	<213>	Artificial	
25	<220>		
	<223>	Identificador molecular	
	<400>	32	
30		acgctcgaca	10
	<210>	33	
	<211>	10	
	<212>	ADN	
35	<213>	Artificial	
	<220>		
	<223>	Identificador molecular	
40	<400>	33	
		agacgcactc	10
	<210>	34	
	<211>	10	
45	<212>	ADN	
	<213>	Artificial	
	<220>		
	<223>	Identificador molecular	
50	<400>	34	
		agcactgtag	10
	<210>	35	
55	<211>	10	
	<212>	ADN	
	<213>	Artificial	
	<220>		
60	<223>	Identificador molecular	
	<400>	35	
		atcagacacg	10

ES 2 676 245 T3

	<210>	36	
	<211>	26	
	<212>	ADN	
5	<213>	Artificial	
	<220>		
	<223>	Cebador inverso	
10	<400>	36	
		ttagatgagc atgatgtaat gagctc	26
	<210>	37	
	<211>	26	
15	<212>	ADN	
	<213>	Artificial	
	<220>		
	<223>	Cebador inverso	
20	<400>	37	
		ttagatgagc atgatgtaat tagctc	26
	<210>	38	
25	<211>	26	
	<212>	ADN	
	<213>	Artificial	
	<220>		
30	<223>	Cebador inverso	
	<400>	38	
		ttagatgagc atgatgtaat cagctc	26
35	<210>	39	
	<211>	26	
	<212>	ADN	
	<213>	Artificial	
40	<220>		
	<223>	Cebador inverso	
	<400>	39	
		ttagatgagc atgatgttat gagctc	26
45	<210>	40	
	<211>	26	
	<212>	ADN	
	<213>	Artificial	
50	<220>		
	<223>	Cebador inverso	
	<400>	40	
55		ttagatgagc atgatgttat tagctc	26
	<210>	41	
	<211>	26	
	<212>	ADN	
60	<213>	Artificial	
	<220>		
	<223>	Cebador inverso	

	<40>	41		
			ttagatgagc atgatgttat cagctc	26
5	<210>	42		
	<211>	26		
	<212>	ADN		
	<213>	Artificial		
10	<220>			
	<223>	Cebador inverso		
	<400>	42		
			ttagatgagc aggatgtaat gagctc	26
15	<210>	43		
	<211>	26		
	<212>	ADN		
	<213>	Artificial		
20	<220>			
	<223>	Cebador inverso		
	<400>	43		
25			ttagatgagc aggatgtaat tagctc	26
	<210>	44		
	<211>	26		
	<212>	ADN		
30	<213>	Artificial		
	<220>			
	<223>	Cebador inverso		
35	<400>	44		
			ttagatgagc aggatgtaat cagctc	26
	<210>	45		
	<211>	26		
40	<212>	ADN		
	<213>	Artificial		
	<220>			
	<223>	Cebador inverso		
45	<400>	45		
			ttagatgagc aggatgttat gagctc	26
	<210>	46		
50	<211>	26		
	<212>	ADN		
	<213>	Artificial		
	<220>			
55	<223>	Cebador inverso		
	<400>	46		
			ttagatgagc aggatgttat tagctc	26
60	<210>	47		
	<211>	26		
	<212>	ADN		
	<213>	Artificial		

	<220>		
	<223>	Cebador inverso	
	<400>	47	
5		ttagatgagc aggatgttat cagctc	26
	<210>	48	
	<211>	26	
10	<212>	ADN	
	<213>	Artificial	
	<220>		
	<223>	Cebador inverso	
15	<400>	48	
		ttagatgagc aagatgtaat gagctc	26
	<210>	49	
	<211>	26	
20	<212>	ADN	
	<213>	Artificial	
	<220>		
	<223>	Cebador inverso	
25	<400>	49	
		ttagatgagc aagatgtaat tagctc	26
	<210>	50	
30	<211>	26	
	<212>	ADN	
	<213>	Artificial	
	<220>		
35	<223>	Cebador inverso	
	<400>	50	
		ttagatgagc aagatgtaat cagctc	26
40	<210>	51	
	<211>	26	
	<212>	ADN	
	<213>	Artificial	
45	<220>		
	<223>	Cebador inverso	
	<400>	51	
		ttagatgagc aagatgttat gagctc	26
50	<210>	52	
	<211>	26	
	<212>	ADN	
	<213>	Artificial	
55	<220>		
	<223>	Cebador inverso	
	<400>	52	
60		ttagatgagc aagatgttat tagctc	26
	<210>	53	
	<211>	26	
	<212>	ADN	

	<213>	Artificial	
	<220>		
	<223>	Cebador inverso	
5	<400>	53	
		ttagatgagc aagatgttat cagctc	26
	<210>	54	
10	<211>	25	
	<212>	ADN	
	<213>	Artificial	
	<220>		
15	<223>	Cebador directo	
	<400>	54	
		catatgatac cagatgcttt gactc	25
	<210>	55	
20	<211>	25	
	<212>	ADN	
	<213>	Artificial	
	<220>		
25	<223>	Cebador directo	
	<400>	55	
		catatgatac cagatgcttt gactc	25
30	<210>	56	
	<211>	25	
	<212>	ADN	
	<213>	Artificial	
35	<220>		
	<223>	Cebador directo	
	<400>	56	
40		catatgatac cagttgcttt gactc	25
	<210>	57	
	<211>	25	
	<212>	ADN	
45	<213>	Artificial	
	<220>		
	<223>	Cebador directo	
50	<400>	57	
		catatgatac cagttgcttt gactc	25
	<210>	58	
	<211>	25	
55	<212>	ADN	
	<213>	Artificial	
	<220>		
	<223>	Cebador directo	
60	<400>	58	
		catatgatac caggtgcttt gactc	25
	<210>	59	

ES 2 676 245 T3

	<211>	25	
	<212>	ADN	
	<213>	Artificial	
5	<220>		
	<223>	Cebador directo	
	<400>	59	
		catatgatac caggtgtttt gactc	25
10	<210>	60	
	<211>	25	
	<212>	ADN	
	<213>	Artificial	
15	<220>		
	<223>	Cebador directo	
	<400>	60	
20		catatgatac cagctgcttt gactc	25
	<210>	61	
	<211>	25	
	<212>	ADN	
25	<213>	Artificial	
	<220>		
	<223>	Cebador directo	
30	<400>	61	
		catatgatac cagctgtttt gactc	25
	<210>	62	
	<211>	25	
35	<212>	ADN	
	<213>	Artificial	
	<220>		
	<223>	Cebador directo	
40	<400>	62	
		catatgatac ccgatgcttt gactc	25
	<210>	63	
45	<211>	25	
	<212>	ADN	
	<213>	Artificial	
	<220>		
50	<223>	Cebador directo	
	<400>	63	
		catatgatac ccgatgtttt gactc	25
55	<210>	64	
	<211>	25	
	<212>	ADN	
	<213>	Artificial	
60	<220>		
	<223>	Cebador directo	
	<400>	64	
		catatgatac ccgttgcttt gactc	25

ES 2 676 245 T3

	<210>	65	
	<211>	25	
	<212>	ADN	
5	<213>	Artificial	
	<220>		
	<223>	Cebador directo	
10	<400>	65	
		catatgatac ccggtgtttt gactc	25
	<210>	66	
	<211>	25	
15	<212>	ADN	
	<213>	Artificial	
	<220>		
	<223>	Cebador directo	
20	<400>	66	
		catatgatac ccggtgcttt gactc	25
	<210>	67	
25	<211>	25	
	<212>	ADN	
	<213>	Artificial	
	<220>		
30	<223>	Cebador directo	
	<400>	67	
		catatgatac ccggtgtttt gactc	25
35	<210>	68	
	<211>	25	
	<212>	ADN	
	<213>	Artificial	
40	<220>		
	<223>	Cebador directo	
	<400>	68	
45		catatgatac ccgctgcttt gactc	25
	<210>	69	
	<211>	25	
	<212>	ADN	
50	<213>	Artificial	
	<220>		
	<223>	Cebador directo	
	<400>	69	
55		catatgatac ccgctgtttt gactc	25
	<210>	70	
	<211>	26	
	<212>	ADN	
60	<213>	Artificial	
	<220>		
	<223>	Cebador inverso	

ES 2 676 245 T3

	<400>	70	gaatatctgg tcatagcatc cgtgaa	26
5	<210>	71		
	<211>	26		
	<212>	ADN		
	<213>	Artificial		
	<220>			
10	<223>	Cebador inverso		
	<400>	71	gaatatctgg tcatagcttc cgtgaa	26
15	<210>	72		
	<211>	26		
	<212>	ADN		
	<213>	Artificial		
20	<220>			
	<223>	Cebador inverso		
	<400>	72	gaatatctgg tcatagcgtc cgtgaa	26
25	<210>	73		
	<211>	26		
	<212>	ADN		
	<213>	Artificial		
30	<220>			
	<223>	Cebador inverso		
	<400>	73	gaatatctgg tcatagcctc cgtgaa	26
35				

REIVINDICACIONES

1. Un kit para detectar una o más variantes de secuencia del virus de la hepatitis C humano que comprende los siguientes conjuntos de cebadores:
- 5
- un primer conjunto de cebadores que comprende un oligonucleótido de fórmula (I) como cebador inverso, y un cebador directo que puede hibridar con la parte 5'-UTR-Núcleo del VHC, en el que el conjunto de cebadores puede generar un fragmento que comprende desde los nucleótidos 146 a 490 de la SEQ ID NO: 2, en el que la fórmula (I) es:
- 10
- $(N)_m-Z$ (I),
- en la que
- 15
- Z es un oligonucleótido que consiste en la SEQ ID NO: 1,
- m es un número entero que varía de 0 a 25 nucleótidos, y
- 20
- N es un nucleótido seleccionado de A, T, C o G;
- un segundo conjunto de cebadores que comprende un oligonucleótido de fórmula (II) como cebador inverso, y un cebador directo que puede hibridar con la parte NS5B del VHC, en el que el conjunto de cebadores puede generar un fragmento que comprende desde el nucleótido 653 al 1106 de la SEQ ID NO: 6, en el que la fórmula (II) es:
- 25
- $(N)_m-F$ (II)
- en la que
- 30
- F es un oligonucleótido que consiste en la SEQ ID NO: 5 (5'- TTNGADGAGCADGATGTWATBAGCTC-3'), en la que N significa adenina (A), citosina (C), guanina (G) o timina (T);
- D significa G, A o T;
- 35
- W significa A o T,
- B significa G, T o C, y
- 40
- m es un número entero que varía de 0 a 25 nucleótidos;
- un tercer conjunto de cebadores que comprende un oligonucleótido que consiste en la SEQ ID NO: 7 como cebador directo, y un cebador inverso que puede hibridar con la parte NS5B del VHC, en el que el conjunto de cebadores puede generar un fragmento que comprende desde los nucleótidos 653 a 1040 de la SEQ ID NO: 6.
- 45
2. Un procedimiento *in vitro* para determinar en una muestra aislada de un sujeto la presencia de una o más variantes de secuencia del virus de la hepatitis C (VHC) humano que comprende:
- 50
- i) transcribir inversamente y amplificar moléculas de ARN del VHC mediante reacción en cadena de la polimerasa con transcripción inversa (RT-PCR) utilizando el primer conjunto de cebadores como se define en la reivindicación 1, y el segundo conjunto de cebadores como se define en la reivindicación 1,
- 55
- ii) amplificar adicionalmente los fragmentos generados en la etapa i) mediante PCR usando el primer conjunto de cebadores como se define en la reivindicación 1; y el tercer conjunto de cebadores como se define en la reivindicación 1, y
- iii) secuenciar los fragmentos amplificados obtenidos en la etapa ii).
3. El kit de la reivindicación 1 o el procedimiento de la reivindicación 2, en el que en la fórmula (I) m es 0, y el cebador inverso del primer conjunto de cebadores consiste en la SEQ ID NO: 1.
- 60
4. El kit de la reivindicación 1 o el procedimiento de la reivindicación 2, en el que el cebador directo del primer conjunto de cebadores consiste en una secuencia de nucleótidos seleccionada de la SEQ ID NO: 3 y la SEQ ID NO: 4.
- 65
5. El kit de la reivindicación 1 o el procedimiento de la reivindicación 2, en el que en la fórmula (II) m es 0 y el

cebador inverso del segundo conjunto de cebadores consiste en la SEQ ID NO: 5.

- 5 6. El kit de la reivindicación 1 o el procedimiento de la reivindicación 2, en el que el cebador directo del segundo conjunto de cebadores consiste en el oligonucleótido de la SEQ ID NO: 7:
- 5'-CNTAYGAYACCMGNTGYTTTGACTC-3' (SEQ ID NO: 7), en la que
- N significa adenina (A), citosina (C), guanina (G) o timina (T);
- 10 Y significa T o C; y
- M significa A o C.
- 15 7. El kit de la reivindicación 1 o el procedimiento de la reivindicación 2, en el que el cebador inverso del tercer conjunto de cebadores consiste en la SEQ ID NO: 8.

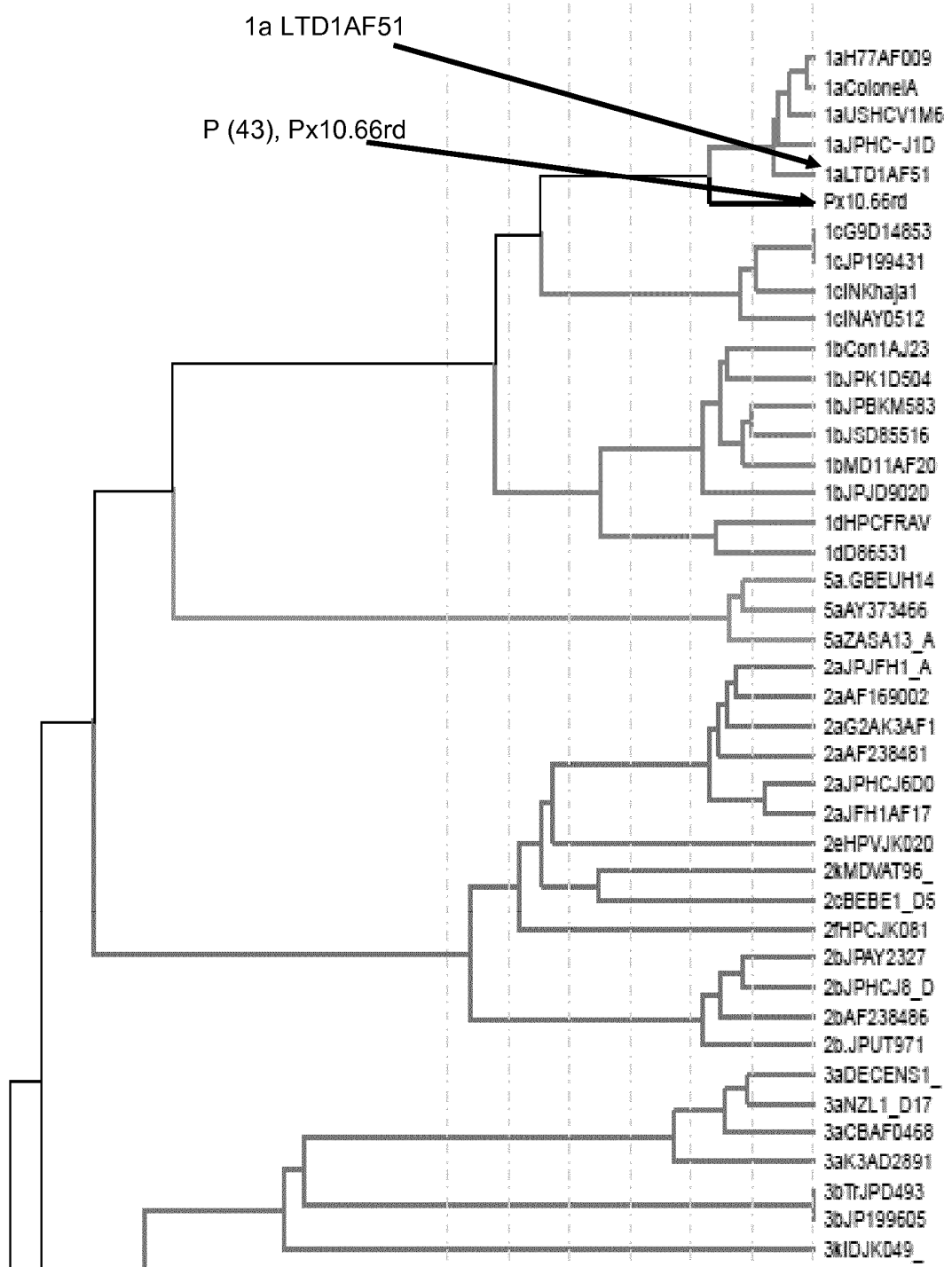


FIG. 1

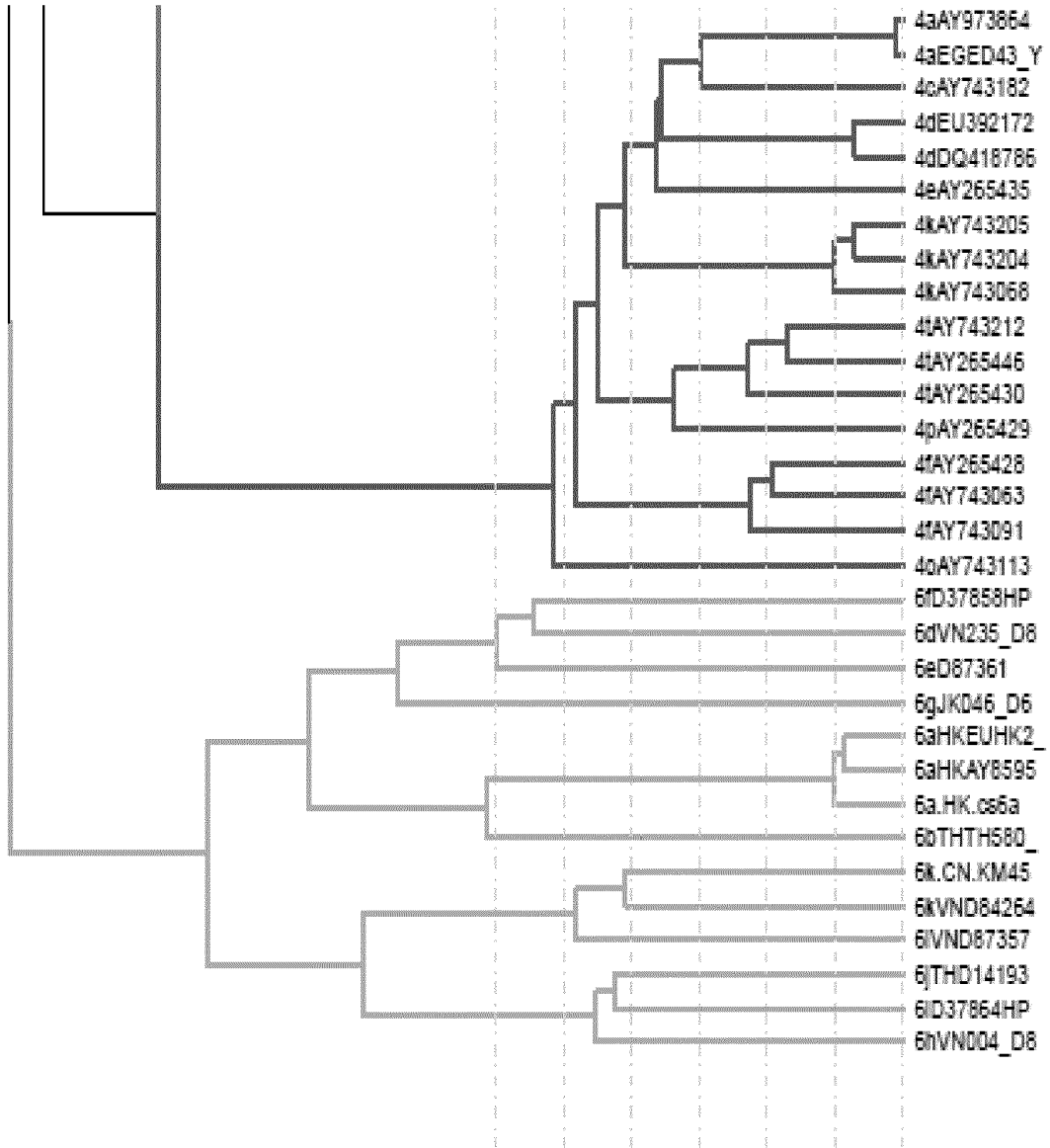


FIG. 1 (Cont.)

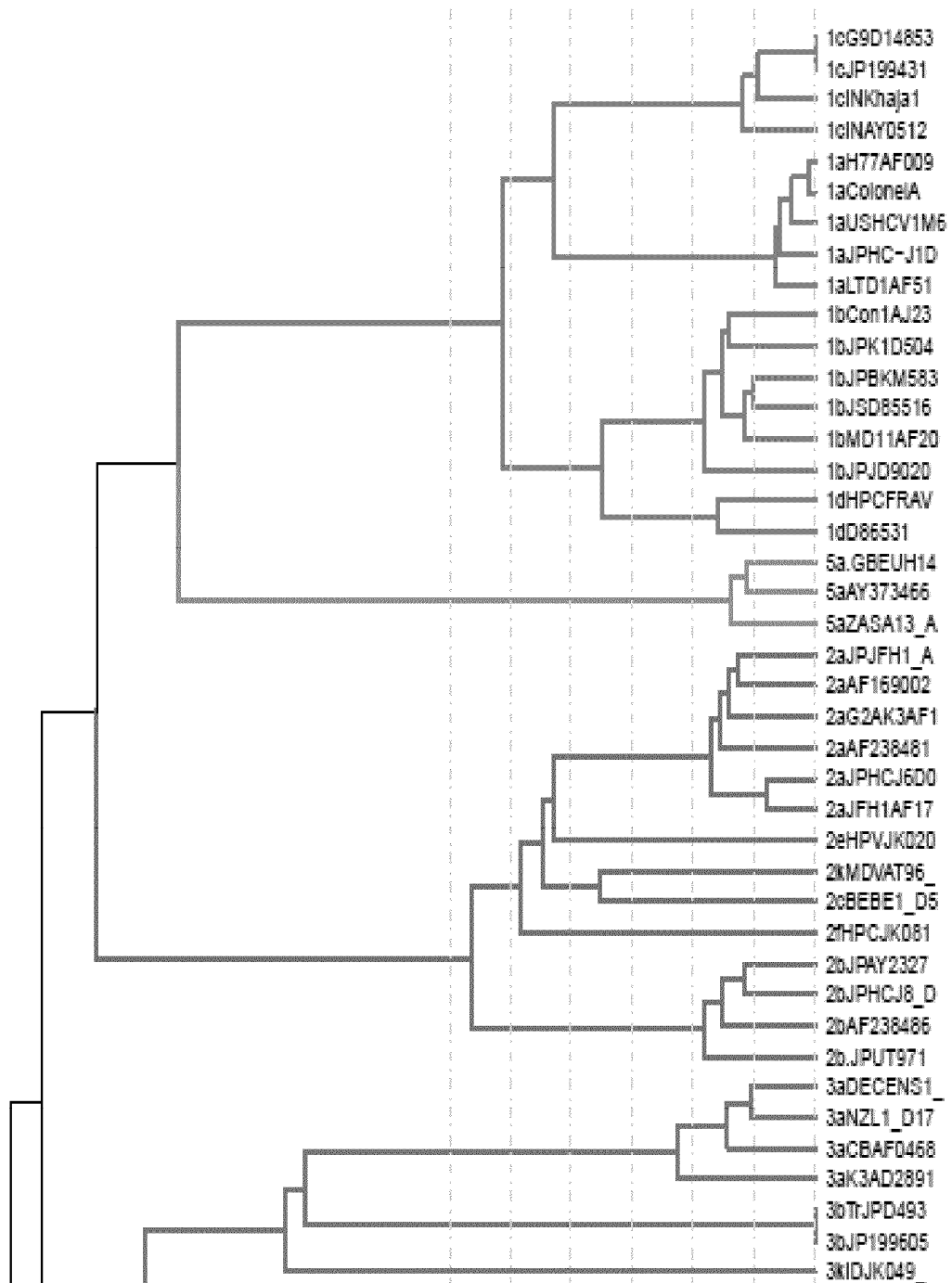


FIG. 2

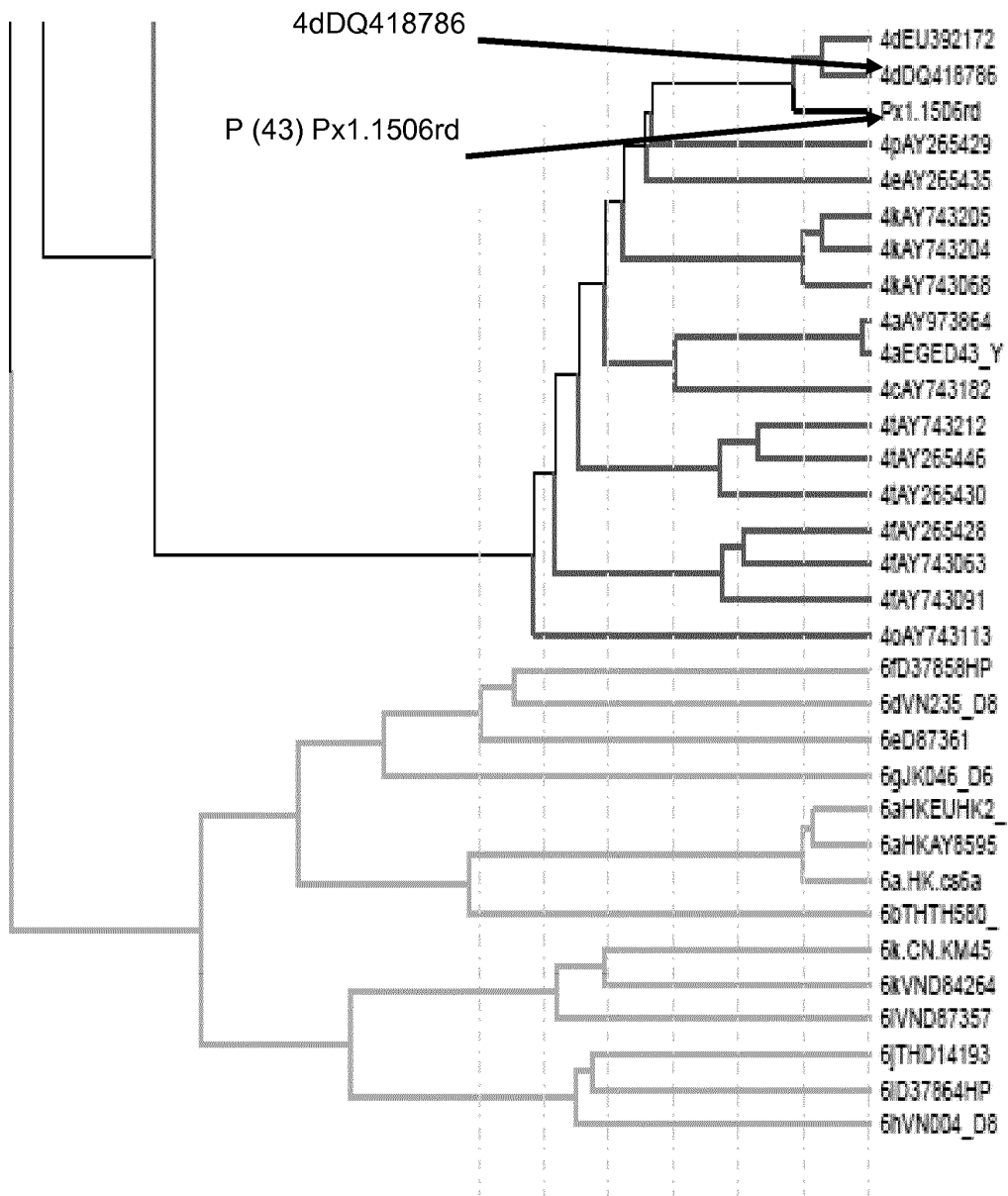


FIG. 2 (Cont.)