(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 297 857**

(51) Int. Cl.:
C12Q 1/68 (2006.01)
C12Q 1/70 (2006.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(86) Número de solicitud europea: **97920354 .4**
(86) Fecha de presentación : **11.04.1997**
(87) Número de publicación de la solicitud: **0904406**
(87) Fecha de publicación de la solicitud: **31.03.1999**

(54) Título: **Ensayo de rastreo de heterodúplex (HTA) para establecer el genotipo de HCV.**

(30) Prioridad: **19.04.1996 US 634797**

(45) Fecha de publicación de la mención BOPI:
01.05.2008

(45) Fecha de la publicación del folleto de la patente:
01.05.2008

(73) Titular/es: **Novartis Vaccines and Diagnostics, Inc.**
4560 Horton Street
Emeryville, California 94608, US

(72) Inventor/es: **Weiner, Amy, J.**

(74) Agente: **Elzaburu Márquez, Alberto**

DESCRIPCIÓN

Ensayo de rastreo de heterodúplex (HTA) para establecer el genotipo de HCV.

Campo de la invención

Esta invención se refiere a genotipar el virus de la hepatitis C (HCV). En particular, esta invención se refiere a cebadores específicos preferiblemente de la región de núcleo y cubierta del HCV y a un método para determinar los genotipos del HCV con un ensayo de movilidad o rastreo del heterodúplex que, a su vez, utiliza cebadores específicos.

Antecedentes de la invención

La hepatitis vírica es conocida por estar causada por cinco virus diferentes conocidos como hepatitis A, B, C, D y E. El HAV es un virus de ARN y no conduce a síntomas clínicos a largo plazo. El HBV es un virus de ADN. El HDV es un virus dependiente que es incapaz de infectar células en ausencia del HBV. El HEV es un virus transmitido por el agua. El HCV se identificó y caracterizó por primera vez como una causa de hepatitis no A no B (NANBH) (Houghton *et al*., pub. EPO nº 388.232 y 318.216). Esto condujo a la descripción de una serie de polipéptidos generales y específicos útiles como reactivos inmunológicos en la identificación del HCV. Véanse, por ejemplo, Choo *et al*. (1989) Science, 244: 359-362; Kuo *et al*., (1989) Science 244: 362-364 y Houghton *et al.*, (1991) Hepatology 14: 381-388.

El HCV es un virus de ARN monocatenario relacionado distantemente con los pestivirus y flavivirus y es el agente causante de la amplia mayoría de las hepatitis asociadas a transfusión y de la mayoría de los casos de hepatitis no A no B adquirida de forma extrahospitalaria en todo el mundo. El genoma del HCV consiste en regiones 5' y 3' no codificantes (NC) que flanquean un marco de lectura abierto (ORF) largo único. Este ORF codifica tres proteínas estructurales en el extremo aminoterminal y seis proteínas no estructurales (NS) en el extremo carboxiterminal. Las proteínas estructurales se representan por las proteínas de nucleocápsida (núcleo, C) y dos glicoproteínas, cubierta 1 (E1) y cubierta 2 (E2). Las proteínas no estructurales se nombran NS2, NS3, NS4a, NS4b, NS5a, NS5b. La 5'NCR es la parte más altamente conservada del genoma del HCV, mientras que la secuencia de las dos proteínas de cubierta (E1 y E2) es altamente variable entre diferentes aislamientos de HCV. El mayor grado de variación se ha observado en una región dentro de E2, ahora denominada habitualmente región hipervariable 1 (HVR1) o E2HV. Existe también una segunda región variable denominada HVR2 en un subconjunto de aislamientos. Típicamente, la heterogeneidad genética del HCV se ha clasificado según las dos denominaciones cuasiespecies y genotipos. Como se utiliza en la presente memoria, el término "cuasiespecie" designa la heterogeneidad genética de la población de HCV dentro de un individuo afectado. Como se utiliza en la presente memoria, los términos "genotipo" y "subtipo" designan la heterogeneidad genómica observada entre diferentes aislamientos de HCV. El análisis de la variación de la secuencia de ácido nucleico del genoma del HCV, una molécula de ARN de cadena positiva de aproximadamente 9,4 kb, sugiere que la variabilidad genética está asociada a importantes implicaciones virológicas y clínicas.

El aislamiento prototípico del HCV se caracterizó en las publicaciones EP nº 318.216 y 388.232. Como se utiliza en la presente memoria, el término "HCV" incluye especies víricas NANBH recién aisladas. El término "HCV-1" designa el virus descrito en las publicaciones anteriormente mencionadas.

Desde la identificación inicial del HCV, se han identificado al menos 6 tipos víricos principales diferentes (genomas completos reseñados) y se han denominado de tipo 1, 2, 3, 4, 5 y 6. Dentro de estos tipos hay numerosos subtipos. El tipo de virus con el que se infecta un paciente puede afectar al pronóstico clínico y responder también a diversos tratamientos. Véase Yoshioke *et al.* (1992) Hepatology 16: 293-299. Considerando que el resultado clínico más grave de la infección por HCV es el carcinoma hepatocelular, sería útil ser capaz de determinar con cuál tipo o tipos de HCV se ha infectado un paciente. Por tanto, es de particular importancia desarrollar un ensayo exacto y fiable de genotipado y subtipado de HCV que, sin requerir secuenciación, pueda proporcionar también la divergencia genética intrasubtipos. Se han propuesto diversas clasificaciones para el genotipado de HCV basándose en el análisis de diferentes regiones, porque el sistema ideal basado en la secuencia nucleotídica que utiliza el genoma vírico completo no es práctico.

Sumario de la invención

La presente invención incluye cebadores y métodos para la caracterización del genotipado de HCV y de la variación intrasubtipos basados en el ensayo de rastreo de heterodúplex (HTA). Las sondas/cebadores preferidos eran monocatenarios derivados de la terminación carboxilo del núcleo y parte de la región E1 del HCV.

El HTA es un método basado en la hibridación de determinación de la relación genética entre dos o más genomas víricos. La base del método es que los productos de ADN relacionados coamplificados a partir de moldes divergentes se reasocian aleatoriamente formando heterodúplex que migran con movilidad reducida en sistemas diseñados para separar moléculas basándose en el tamaño, tales como los geles de poliacrilamida neutra. El HTA se utilizó originalmente para genotipar VIH-1 y para seguir la evolución *in vivo* de VIH-1 en pacientes y poblaciones. Véanse, por ejemplo, Delwart *et al.*, (1993) Science 262: 1757-1261 y Delwart *et al.*, (1994) J. Virol. 68: 6772-6883.

Wilson *et al*., J. Gen. Virol. (1995) 76: 1763-1771, dan a conocer el uso de análisis de desplazamiento en gel de heterodúplex (GSA) de secuencias de HVR1 amplificadas directamente a partir de sueros de pacientes para caracterizar cuasiespecies de HCV.

La invención proporciona un oligonucleótido consistente en una cualquiera de las secuencias de SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, 8 y 9.

La invención proporciona también un par de cebadores de PCR en los que el cebador con sentido consiste en la SEQ ID NO: 1 y el cebador antisentido se selecciona del grupo consistente en las SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4 y SEQ ID NO: 5.

La invención proporciona también un par de cebadores de PCR en los que el cebador antisentido consiste en la SEQ ID NO: 6 y el cebador con sentido se selecciona del grupo consistente en las SEQ ID NO: 7, SEQ ID NO: 8 y SEQ ID NO: 9.

La invención proporciona también un método de determinación del genotipo del HCV de una cepa de HCV, comprendiendo dicho método las etapas de:

(a) someter dicha cepa de HCV a una o más etapas de PCR, en el que las una o más etapas de PCR utilizan una sonda con sentido del núcleo o de la región E1 del genoma del HCV y una sonda antisentido del núcleo o de la región E1 del genoma del HCV;

(b) formar un heterodúplex mediante la desnaturalización de las mezclas de reasociación del producto amplificado obtenido en la etapa (a) con fragmentos de ADN o ARN de un genotipo de HCV conocido;

(c) comparar la movilidad de dicho heterodúplex en un sistema que separa por tamaño con la movilidad de un homodúplex de fragmentos de ADN o ARN de genotipo conocido para determinar el genotipo de la cepa del HCV.

La invención proporciona también un método para predecir la respuesta a la terapia farmacológica de una cepa de HCV a partir de un paciente infectado con dicha cepa de HCV, comprendiendo dicho método determinar la sensibilidad de genotipos de HCV conocidos frente a dicha terapia farmacológica, determinar el genotipo de HCV de dicha cepa de HCV mediante el método de la invención, y comparar dicho genotipo de HCV de dicha cepa antes de dicha terapia farmacológica con dicha sensibilidad de genotipos de HCV conocidos frente a dicha terapia farmacológica.

La invención proporciona también un método para predecir la respuesta a una vacuna terapéutica de una cepa de HCV a partir de un paciente infectado con dicha cepa de HCV, comprendiendo dicho método determinar la sensibilidad de genotipos de HCV conocidos frente a dicha vacuna terapéutica, determinar el genotipo de HCV de dicha cepa de HCV mediante el método de la invención, y comparar dicho genotipo de HCV de dicha cepa antes de la administración de dicha vacuna terapéutica con dicha sensibilidad de genotipos de HCV conocidos frente a dicha vacuna terapéutica.

La invención proporciona también un método para predecir la idoneidad de una composición de vacuna profiláctica para una población de muestra dada, comprendiendo dicho método determinar el genotipo de dicha vacuna profiláctica, determinar la predominancia de genotipos de HCV conocidos en dicha población de muestra mediante el método de la invención, y comparar dicho genotipo de HCV de dicha cepa de vacuna profiláctica con el genotipo predominante determinado antes de la administración de dicha vacuna profiláctica a dicha muestra de población.

Es un aspecto de la invención un método para genotipar el HCV que comprende las etapas de desnaturalizar y reasociar cadenas de ADN o ARN parcialmente complementarias y detectar la variación de secuencia observando la movilidad electroforética de los heterodúplex de ADN en un sistema diseñado para separar moléculas basándose en el tamaño, tal como siguiendo la electroforesis a través de un gel de poliacrilamida o MDE.

Otro aspecto de la invención se refiere a las sondas utilizadas en el genotipado, que se seleccionaron a partir del núcleo y la región E1 del genoma de HCV.

Otro aspecto de la invención se refiere a un método de predicción de la respuesta a una terapia farmacológica de un paciente infectado con una cepa de HCV determinando la sensibilidad de diferentes genotipos conocidos frente a la terapia farmacológica, determinando el genotipo de la cepa de HCV que infecta al paciente y comparando el genotipo con su sensibilidad frente a la terapia farmacológica para predecir la respuesta del paciente a la terapia farmacológica.

Otro aspecto de la invención se refiere a vacunas terapéuticas y a predecir qué vacuna terapéutica debería utilizarse, determinando el genotipo de un paciente infectado con una cepa de HCV y administrando una vacuna terapéutica del mismo genotipo.

Otro aspecto de la invención se refiere a vacunas profilácticas y a predecir qué vacuna debería administrarse a una cierta muestra de población, determinando los genotipos prevalentes en una muestra similar, y administrando vacunas profilácticas de un genotipo que es probable que sea el genotipo prevalente a la muestra de población.

Otro aspecto de la invención se refiere a la capacidad de descubrir nuevos genotipos de HCV utilizando el método de la invención.

Breve descripción de las figuras

Las Figuras 1A-1E son autorradiogramas que muestran homodúplex y heterodúplex de las muestras para tipar con las sondas de genotipos conocidos (las sondas mc son de los genotipos 1a, 1b, 2a, 2b, 3a en las Figs. 1A-1E, respectivamente, carril al extremo izquierdo del gel MDE). El homodúplex (h) (sonda mc del producto de PCR-TI bicatenario de genotipo conocido del que se derivó) se muestra adyacente a la sonda. Los heterodúplex de los productos de PCR-TI de los 15 pacientes de diálisis (nº 1, 2, 3, 4, 7, 18, 20, 22, 23, 24, 26, 28, 30, 33, 35) hibridados con la sonda mc se denominan encima del carril apropiado en cada figura.

Las Figuras 2A-2C son dendogramas, concretamente, árboles filogenéticos que muestran la relación de cada secuencia nucleotídica de E1 parcial, formados comparando secuencias de E1 parciales obtenidas secuenciando supuestos aislamientos de tipo 1 (nt 625-930), de tipo 2 (nt 583-915) o de tipo 3 (nt 558-834) a partir de los pacientes de diálisis descritos en la presente memoria con secuencias de genotipo publicado para el tipo 1a (HCV-1) (Choo *et al.*, PNAS (1991), 88: 2451-2455, todas las denominaciones de nucleótidos "nt" según este artículo), 1b (HCV-J) (Kato *et al.*, PNAS (1990) 87: 9524-2528), 2a (HC-J6) (Okamoto *et al.*, Virol. (1992) 188: 331-341), 2b (HC-J8) (Okamoto *et al.*, Virol. (1992) 188: 331-341); 2c (Bukh *et al.*, PNAS (1993) 90: 8234-8239) y 3a (NZL-1) (Sakamoto *et al.*) J. Gen. Virol. (1994) 75: 1761-1768 en la misma región del genoma.

La Figura 2D es un dendograma, árbol filogenético, formado comparando secuencias de 5'UTR parciales de los aislamientos 23, 30 y 33 obtenidas mediante secuenciación directa con secuencias de genotipo publicado de tipo 1, 2 y 3 (nt-274 a -81) para la misma región del genoma.

Las Figuras 3A-3D muestran las secuencias nucleotídicas para los dendogramas exhibidos en las Figuras 2A-2D.

Descripción de la invención

La práctica de la presente invención empleará, a menos que se indique otra cosa, técnicas convencionales de biología molecular, microbiología, ADN recombinante, síntesis de polipéptido y ácido nucleico e inmunología, que están dentro de las habilidades de la técnica. Dichas técnicas se explican completamente en la bibliografía. Véanse, por ejemplo, Sambrook *et al.*, "Molecular Cloning: A Laboratory Manual", 2ª edición (1989); "DNA Cloning", volúmenes I y II (D.N. Glover ed. 1985); "Oligonucleotide Synthesis" (M.J. Gait ed., 1984); "Nucleic Acid Hybridization" (B.D. Hames & S.J. Higgins ed., 1984); "Transcription and Translation" (B.D. Hames & S.J. Higgins ed., 1984); "Animal Cell Culture" (R.I. Freshney ed., 1986); "Immobilized Cells and Enzymes" (IRL Press, 1986); B. Perbal, "A Practical Guide to Molecular Cloning" (1984); la serie "Methods in Enzymology" (Academic Press, Inc.); "Gene Transfer Vectors for Mammalian Cells" (J.H. Miller y M.P. Calos, ed. 1987, Cold Spring Harbor Laboratory); Methods in Enzymology vol. 154 y vol. 155 (Wu y Grossman y Wu, ed., respectivamente); Mayer y Walker ed. (1987), "Immunochemical Methods in Cell and Molecular Biology", (Academic Press, Londres); Scopes (1987), "Protein Purification: Principles and Practice", 2º edición (Springer-Verlag, N.Y.) y "Handbook of Experimental Immunology", volúmenes I-IV (D.M. Weir y C.C. Blackwell ed., 1986).

Se utilizan las abreviaturas estándar para nucleótidos y aminoácidos en esta memoria descriptiva.

El término "polinucleótido recombinante" como se utiliza en la presente memoria quiere pretende indicar un polinucleótido de origen genómico, de ADNc, semisintético o sintético que, en virtud de su origen o manipulación:

(1) no está asociado con todo o con una porción de un polinucleótido con el que está asociado en la naturaleza, (2) está ligado con un polinucleótido distinto al que está ligado en la naturaleza, o (3) no aparece en la naturaleza.

El término "polinucleótido" como se utiliza en la presente memoria designa una forma polimérica de nucleótidos de cualquier longitud, ribonucleótidos o desoxirribonucleótidos. Este término designa sólo la estructura primaria de la molécula. Por tanto, este término incluye ADN y ARN bi- y monocatenarios. Incluye también tipos conocidos de modificaciones, por ejemplo, marcadores que son conocidos en la técnica, metilación, "caperuzas", sustitución de uno o más de los nucleótidos de origen natural por un análogo, modificaciones internucleotídicas tales como, por ejemplo, aquellas con ligamientos no cargados (por ejemplo, metilfosfonatos, fosfotriésteres, fosfoamidatos, carbamatos, etc.) y con ligamientos cargados (por ejemplo, fosforotioatos, fosforoditioatos, etc.), aquellas que contienen restos pendientes tales como, por ejemplo, proteínas (incluyendo, por ejemplo, nucleasas, toxinas, anticuerpos, péptidos señal, poli-L-lisina, etc.), aquellas con intercalantes (por ejemplo, acridina, psoraleno, etc.), aquellas que contienen quelantes (por ejemplo, metales, metales radiactivos, boro, metales oxidantes, etc.), aquellas que contienen alquilantes, aquellas con ligamientos modificados (por ejemplo, ácidos nucleicos alfa-anoméricos, etc.), así como formas modificadas del polinucleótido.

Por "PCR" se quiere indicar en la presente memoria la técnica de reacción en cadena de la polimerasa (PCR), dada a conocer por Mullis en las patentes de EE.UU. nº 4.683.195 (Mullis *et al.*) y 4.683.202. En la técnica de PCR, se preparan cebadores oligonucleotídicos cortos que coinciden con los extremos opuestos de una secuencia deseada. La secuencia entre los cebadores no necesita conocerse. Se extrae una muestra de ADN (o ARN) y se desnaturaliza (preferiblemente por calor). Después, se añaden cebadores oligonucleotídicos en exceso molar, junto con dNTP y una polimerasa (preferiblemente polimerasa Taq, que es estable al calor). Se replica el ADN, después se desnaturaliza de nuevo. Esto da como resultado dos "productos largos" que empiezan con los cebadores respectivos, y las dos cadenas

originales (por molécula de ADN dúplex). Se devuelve después la mezcla de reacción a condiciones de polimerización (por ejemplo, reduciendo la temperatura, inactivando un agente desnaturalizante o añadiendo más polimerasa), y se inicia un segundo ciclo. El segundo ciclo proporciona las dos cadenas originales, los dos productos largos del ciclo 1, dos productos largos nuevos (replicados a partir de las cadenas originales), y dos "productos cortos" replicados a partir de los productos largos. Los productos cortos tienen la secuencia de la secuencia diana (con sentido o antisentido) con un cebador en cada extremo. En cada ciclo adicional, se producen dos productos largos adicionales y un número de productos cortos igual al número de productos largos y cortos restantes al final del ciclo anterior. Por tanto, el número de productos cortos crece exponencialmente con cada ciclo. Esta amplificación de una secuencia de analito específica permite la detección de cantidades extremadamente pequeñas de ADN.

El término "3SR" como se utiliza en la presente memoria designa un método de amplificación de ácido nucleico diana también conocido como sistema de "replicación de secuencia autosostenida", como se describe en la publicación de patente europea nº 373.960 (publicada el 20 de junio de 1990).

El término "LCR" como se utiliza en la presente memoria designa un método de amplificación de ácido nucleico diana también conocido como "reacción en cadena de la ligasa", como se describe por Barany, Proc. Natl. Acad. Sci. (USA) (1991) 88: 189-193.

Un "marco de lectura abierto" (ORF) es una región de una secuencia polinucleotídica que codifica un polipéptido; esta región puede representar una porción de una secuencia de codificación o una secuencia de codificación total.

Una "secuencia de codificación" es una secuencia polinucleotídica que se traduce en un polipéptido, habitualmente a través de ARNm, cuando se dispone bajo el control de secuencias reguladoras apropiadas. Los límites de la secuencia de codificación están determinados por un codón de iniciación de la traducción en la terminación 5' y un codón de terminación de la traducción en la terminación 3'. Una secuencia de codificación puede incluir, pero sin limitación, ADNc y secuencias polinucleotídicas recombinantes.

Como se utiliza en la presente memoria, "polipéptido" designa un polímero de aminoácidos y no designa una longitud específica del producto; por tanto, péptidos, oligopéptidos y proteínas están incluidos dentro de la definición de polipéptido. Este término no designa tampoco ni excluye modificaciones post-expresión del polipéptido, por ejemplo, glucosilaciones, acetilaciones, fosforilaciones y similares. Se incluyen dentro de la definición, por ejemplo, polipéptidos que contienen uno o más análogos de un aminoácido (incluyendo, por ejemplo, aminoácidos no naturales, etc.), polipéptidos con ligamientos sustituidos, así como otras modificaciones conocidas en la técnica, tanto de origen natural como de origen no natural.

Un polipéptido o secuencia de aminoácidos "derivado de" una secuencia de ácido nucleico denominada designa un polipéptido que tiene una secuencia de aminoácidos idéntica a la de un polipéptido codificado en la secuencia, o una porción del mismo en la que la porción consiste en al menos 3-5 aminoácidos, y más preferiblemente al menos 8-10 aminoácidos, y aún más preferiblemente al menos 11-15 aminoácidos, o que es inmunológicamente identificable con un polipéptido codificado en la secuencia. Esta terminología incluye también un polipéptido expresado a partir de una secuencia de ácido nucleico denominada.

La proteína puede utilizarse para producir anticuerpos, monoclonales o policlonales, específicos de la proteína. Los métodos para producir estos anticuerpos son conocidos en la técnica.

"Células anfitrionas recombinantes", "células anfitrionas", "células", "cultivos celulares" y otros de dichos términos indican, por ejemplo, microorganismos, células de insecto y células de mamífero que pueden utilizarse, o se han utilizado, como receptores para vector recombinante u otro ADN de transferencia, e incluyen la progenie de la célula original que se ha transformado. Se entiende que la progenie de una célula parental individual puede no ser necesariamente completamente idéntica en morfología o en complementariedad con el ADN genómico o ADN total al progenitor original, debido a mutación natural, accidental o deliberada. Los ejemplos de células anfitrionas de mamífero incluyen células de ovario de hámster chino (CHO) y de riñón de mono (COS).

Por "ADNc" se quiere indicar una secuencia de ARNm complementario que hibrida con una cadena complementaria de ARNm.

Por "purificado" y "aislado" se quiere indicar, cuando se designa un polipéptido o secuencia nucleotídica, que la molécula indicada está presente sustancialmente en ausencia de otras macromoléculas biológicas del mismo tipo. El término "purificado" como se utiliza en la presente memoria significa preferiblemente al menos un 75% en peso, más preferiblemente al menos un 85% en peso, más preferiblemente aún al menos un 95% en peso, y lo más preferiblemente al menos un 98% en peso, de macromoléculas biológicas del mismo tipo presentes (pero agua, tampones y otras moléculas pequeñas, especialmente moléculas que tienen un peso molecular menor de 1.000, pueden estar presentes).

Por "portador farmacéuticamente aceptable" se quiere indicar cualquier portador farmacéutico que no induzca por sí mismo la producción de anticuerpos dañinos para el individuo que recibe la composición. Los portadores adecuados son típicamente macromoléculas grandes metabolizadas lentamente tales como proteínas, polisacáridos, poli(ácidos lácticos), poli(ácidos glicólicos), aminoácidos poliméricos, copolímeros de aminoácidos y partículas víricas inactivas. Dichos portadores son bien conocidos por los expertos en la técnica.

Las composiciones terapéuticas contendrán típicamente vehículos farmacéuticamente aceptables tales como agua, disolución salina, glicerol, etanol, etc. Adicionalmente, pueden estar presentes en dichos vehículos sustancias auxiliares tales como agentes humectantes o emulsionantes, sustancias tamponadoras del pH y similares.

Típicamente, las composiciones terapéuticas se preparan en forma de inyectables, como disoluciones líquidas o suspensiones; pueden prepararse también formas sólidas adecuadas para disolución o suspensión en vehículos líquidos antes de la inyección. La preparación puede ser también emulsionada o encapsulada en liposomas para un efecto coadyuvante potenciado.

Las pruebas indican que diferentes genotipos del HCV pueden tener diferentes patogenicidades, así como distintas distribuciones geográficas, y pueden desencadenar perfiles serológicos parcialmente diferentes en pacientes infectados. Véase Cammarota *et al.*, J. Clin. Microb. (1995) 33: 2781-2784. La invención incluye métodos para la detección del HCV y la identificación de la infección por diferentes tipos del HCV. La invención incluye genotipar el HCV, el potencial de descubrir un nuevo genotipo del HCV y evaluar en las poblaciones víricas la capacidad de predecir la respuesta a terapia farmacológica. La invención incluye también sondas para uso en el genotipado del HCV.

Los métodos para genotipar el HCV incluyen, pero sin limitación, un ensayo de rastreo o movilidad de heterodúplex que utiliza sondas/cebadores de la región del núcleo/E1 del genoma del HCV. Las diferencias antigénicas documentadas entre genotipos del HCV tendrían utilidad no sólo en el examen de donantes de sangre y para predecir la respuesta al tratamiento con IFN, sino también para la composición denominada de vacunas candidatas para el HCV en diferentes países, la elección de vacunas terapéuticas, así como la identificación de nuevos genotipos. Se han propuesto otros métodos para identificar los genotipos principales que infectan a las poblaciones basándose en el análisis de diferentes regiones del genoma, tales como RFLP. Véase Davidson *et al.*, J. Gen. Virol. (1995) 76: 1197-1204 para la discusión del genotipado de HCV utilizando RFCP de secuencias coamplificadas a partir de la región 5' no codificante (NCR).

Los métodos de genotipado basados en ácido nucleico conocidos requieren (1) cebadores de PCR-TI (PCR-transcriptasa inversa) específicos de un subtipo (véanse Okamoto (1992), J. Gen. Virol. 73: 673-679), Patente de EE.UU. 5.427.909; (2) sondas específicas (G. Marteen *et al.*, "Line probe assay"); (3) polimorfismo de sitio de restricción (una función de la secuencia nucleotídica (nt)) o (4) secuencias directas para determinar el genotipo. El análisis de la secuencia 5'NC con RFLP es fácil de realizar, pero no predice exactamente todos los genotipos del HCV, y algunos subtipos pueden clasificarse erróneamente. Por ejemplo, el cambio de secuencia entre 1a y 1b reconocido por la enzima de restricción no es absoluto, y secuencias distintas de 1a y 1b y 2a y 2b se clasifican erróneamente. Por ejemplo, el tipo 1c aparecería como tipo 1a, el tipo 2c como tipo 2a o 2b. Véase Cammarota *et al.*, J. Clin. Microb. (1995) 33: 2781-2784. Por esta razón, el RFLP no es capaz de detectar especies "de escape", nuevas especies divergentes o tendencias epidemiológicas. Es probable que un método de tipado como RFLP tenga que modificarse continuamente para acomodarse a la información rápidamente creciente recogida sobre la heterogeneidad de secuencia del HCV.

Como se menciona anteriormente, cuando se utilizan los métodos de genotipado basados en ácido nucleico, se obtiene un resultado de un tipo o subtipo o negativo, que es un resultado "no tipable". Véase, por ejemplo, Cammarota *et al.*, J. Clin. Microb. (1995) 33: 2781-2784, los aislamientos que permanecieron no tipados mediante PCR específica de genotipo se clasificaron como subtipo 2c basándose en el análisis de secuencia de amplificaciones por PCR obtenidas a partir de genes del núcleo y NS5. Este problema se evita utilizando la invención actualmente reivindicada para determinar genotipos de HCV eligiendo cebadores de PCR-TI en la terminación C o núcleo/segundo tercio central de E1. Además, el subtipo del aislamiento puede determinarse exactamente utilizando el genotipado de HCV de la presente invención y pueden detectarse aislamientos, incluso aquellos menos de aproximadamente 30% divergentes, posibilitando la caracterización de nuevos subtipos sin secuenciación.

Ensayo de rastreo o movilidad de heterodúplex

El método de determinación del genotipo del HCV de la presente invención utiliza variantes menores en cuasiespecies complejas. Una de dichas técnicas es el ensayo de rastreo de heterodúplex (HTA). El HTA, bien conocido en la técnica para uso con el VIH (véanse, por ejemplo, Delwart *et al.*, J. Virol. (1994) 68: 6672-6683; Delwart *et al.*, Science (1993) 262: 1257-1261; Delwart *et al.*, "PCR Methods and Applications" 4: S202-S216 (1995) Cold Spring Harbor; y Delwart *et al.*, "Heteroduplex Mobility Analysis HIV-1 *env* Subtyping Kit Protocol version 3", surgió de la observación de que cuando se amplificaban secuencias mediante PCR anidada a partir de células mononucleares de sangre periférica de individuos infectados, los productos de ADN relacionados coamplificados a partir de moldes divergentes podían reasociarse aleatoriamente formando heterodúplex que migran con movilidad reducida en geles de poliacrilamida neutra. Utilizando estas técnicas, pueden establecerse las relaciones genéticas entre moléculas molde de ADN vírico múltiples.

El HTA en particular utiliza un primer producto de PCR como sonda marcada, que puede ser radiactiva, que se mezcla con un exceso (promotor) de un producto de PCR no marcado de una fuente diferente, concretamente, la fuente de la que se desea el tipado. Las secuencias de sonda se promueven entonces completamente a heterodúplex con el promotor, y se separan basándose en el tamaño. Un autorradiograma, por ejemplo del gel de poliacrilamida resultante, revela sólo estos heterodúplex y proporciona una representación visual de la relación entre las dos poblaciones de virus en estudio. El hecho de que los heterodúplex migren con distintas movilidades indica que la composición específica de cadena de nucleótidos apareados erróneamente y desapareados afecta a su movilidad.

Se utiliza entonces una ecuación exponencial descrita en Delwart *et al.* para describir un ajuste de curva de los datos experimentales a partir del análisis pareado de genes de secuencia conocida. En la presente invención, se utiliza la ecuación para estimar la distancia genética entre los genotipos conocidos de las sondas y los genotipos desconocidos de las muestras de paciente.

Cebadores para uso en el HTA

Se ha determinado que la región E1 o de núcleo podría ser la mejor región para estudiar la heterogeneidad del HCV, por tanto, la región E1 se convirtió en la elección para cebadores de la presente invención. El uso de la secuencia de E1 parcial, la región más heterogénea del genoma para la presente invención, así como de un fragmento más largo, concretamente de 400 nt, aunque podría haber sido tan largo como de 1.000 nt, posibilitó el diseño de sondas que no hibridan cruzadamente entre subtipos/tipos y por tanto permiten un genotipado exacto. Flanqueando la región heterogénea, se identificaron secuencias de nt conservadas para cebadores con sentido y antisentido. Preferiblemente, se utilizaron una combinación de cebadores con sentido universales y antisentido específicos de tipo para la primera ronda de PCR, y cebadores antisentido universales y con sentido específicos de tipo para la segunda ronda. La PCR no necesita ser de dos rondas, y los cebadores no se limitan a la combinación anteriormente descrita. Sin embargo, la combinación preferida posibilitaba la preparación de sondas monocatenarias y minimizaba el número de combinaciones de cebador de PCR.

Las sondas preferidas son secuencias en las regiones de núcleo y E1 de las que se han publicado las secuencias para un amplio intervalo de genotipos y agrupado en al menos 12 genotipos y subtipos distintos: I/1a, II/1b, III/2a, IV/2b, 2c, 3a, 4a, 4b, 4c, 4d, 5a, 6a. Las identidades de la secuencia nucleotídica del gen E1 entre aislamientos de HCV del mismo genotipo están en el intervalo de 88,0% a 99,1%, mientras que las de aislamientos de HCV de genotipos diferentes están en el intervalo de 53,5 a 78,6%. El grado de variación para una buena discriminación de heterodúplex en geles de poliacrilamida neutra está cómodamente en el intervalo de 3-20%, así que es probable que los moldes divergentes se reasocien formando un heterodúplex si son del mismo subtipo. Por esta razón, se utilizó una sonda de ADN monocatenaria marcada con ^{32}P de modo que, si la formación del heterodúplex es imposible, la sonda de ADNmc pueda probablemente no reasociarse y formar una banda de homodúplex. Sin secuenciación directa, la presente invención puede proporcionar rápidamente no sólo una identificación cierta de los subtipos, sino también de las relaciones genéticas dentro de los mismos subtipos. Por ejemplo, los genotipos utilizados, concretamente (1a, 1b, 2a, 2b, 3a) no mostraron superposición entre subtipos diferentes.

Además, puesto que pueden visualizarse en el gel aislamientos aproximadamente un 30% divergentes, pueden visualizarse nuevos subtipos, podría caracterizarse la distribución de los aislamientos en una población y pueden seguirse aislamientos de poblaciones o individuos en poblaciones o en individuos en estudios epidemiológicos.

Kits de genotipado de HCV

Un kit para determinar el genotipo del HCV está dentro del alcance de esta invención. Como se describe para el VIH en Delwart *et al.*, "Heteroduplex Mobility Analysis HIV *env* Subtyping Kit Protocolo Version 3", dicho kit incluiría cebadores específicos. Los cebadores preferidos son de la región de núcleo y E1 del genoma del HCV. Si se desean dos etapas de PCR, los cebadores de la primera ronda podrían incluir, por ejemplo, una sonda con sentido universal, preferiblemente localizada en la región de núcleo/E1 del genoma de HCV. Uno de dichos cebadores universales está localizado en los nucleótidos 508 a 529 del HCV-1 y se muestra en la Tabla 1. Acoplado al cebador universal, podría haber un cebador antisentido específico de tipo localizado también preferiblemente en la región de núcleo/E1 del genoma de HCV. Son ejemplos de estos cebadores de los nucleótidos 1032 a 1012 para el tipo 1, tipo 2a, tipo 2b y tipo 3a de los genomas de HCV, y se muestran también en la Tabla 1.

Si se desea una segunda ronda de PCR, los cebadores de segunda ronda serían igualmente de la región de núcleo/E1 del genoma de HCV. Los cebadores de segunda ronda preferidos podrían incluir un cebador antisentido universal de los nucleótidos 978 a 958 del genoma del HCV-1, este cebador se muestra en la Tabla 1. Además, los cebadores de segunda ronda podrían incluir un cebador con sentido específico de tipo de la región de núcleo/E1. Los cebadores con sentido específicos de tipo de segunda ronda preferidos son de los nucleótidos 536 a 557 de los genomas de HCV de tipo 1, tipo 2 o tipo 3, y se muestran en la Tabla 1.

La primera o segunda ronda de cebadores pueden ser suficientes para amplificar el ARN vírico, sin utilizar una segunda ronda de PCR si la concentración de los virus es suficientemente alta, concretamente, no se requiere necesariamente PCR anidada, que se requiere en productos de PCR a un exceso de 100x de la sonda.

Un kit de genotipado de HCV de la presente invención incluiría también referencias de subtipo que pueden cambiar a medida que se descubren nuevos subtipos y se evalúan para uso en el kit. Se recomienda el uso de más de una referencia de un subtipo dado porque la comparación con una sola referencia no siempre proporciona un resultado no ambiguo.

La discusión anterior y los ejemplos siguientes sólo ilustran la invención, los expertos en la técnica apreciarán que la invención puede ejecutarse de otros modos, y que la invención se define únicamente con referencia a las reivindicaciones.

Ejemplo 1

Muestras de paciente

Se estudiaron 35 pacientes hemodializados que experimentan hemodiálisis regularmente: 20 hombres (57%) y 15 mujeres (43%) con una edad media de 64,8 ± 13 años. Se recogieron muestras de suero en agosto de 1995, se dividieron en alícuotas y se almacenaron a -80ºC. 26 pacientes eran positivos por ELISA anti-HCV y 9 negativos por ELISA anti-HCV. 25 de los 26 positivos por ELISA eran también positivos de RIBA III, mientras que 1 era indeterminado. Los 9 negativos por ELISA eran todos negativos de RIBA III. 15 pacientes eran positivos por PCR de 5'UTR y E1 de ARN de HCV. Mediante secuenciación directa de 15 productos de 5'NCR, 5 pacientes resultaron de tipo 1, 3 pacientes de tipo 2 y 7 pacientes de tipo 3.

Ejemplo 2

ADNc y PCR

Se extrajo ARN de HCV al menos en dos momentos diferentes utilizando un reactivo Stratagene de un kit de aislamiento de ARN de Stratagene (método de Chomezynsky y Sacchi).

Se extrajo ARN de 20 μl de plasma, que se transcribió inversamente en 25 μl de una mezcla de ADNc (kit de síntesis de ADNc BRL, 8085SB) utilizando 100 pmol de cebadores de PCR. Se hirvió la mezcla de ADNc durante 5 minutos, se enfrió rápidamente en hielo y se añadieron a la PCR los reactivos de ADNc con concentraciones finales según las instrucciones del kit PCR de Perkin (N801-0055). Se realizaron 40 ciclos de PCR (94ºC durante 10 segundos, 55ºC durante 30 segundos y 72ºC durante 30 segundos). Se añadieron 10 μl de la primera mezcla de reacción PCR a una segunda mezcla de reacción PCR que contenía cebadores de PCR anidados y se amplificó durante 40 ciclos como se indica anteriormente.

Se utilizó la primera extracción para la reacción de PCR anidada con cebadores específicos de 5'NCR como se describe anteriormente en Shimizu *et al.*, PNAS (1992) 5477-5481, se secuenció directamente este producto y se utilizó para RFLP. Se utilizó ARN de la misma extracción para HTA utilizando cebadores de núcleo/E1. Se realizó una segunda extracción de ARN para RFLP y/o HTA para confirmar los resultados. Los cebadores utilizados para el HTA se enumeran en la Tabla 1. Los pares de cebadores anidados de PCR utilizados para obtener estos productos de E1 eran diferentes para los tipos 1, 2a, 2b y 3a. La sonda con sentido universal para la primera ronda de amplificación corresponde a los nt 5'-3' 508-529, aminoácidos 170-176, de Choo *et al.*, PNAS 1991, mientras que el cebador antisentido universal para la segunda ronda de amplificación corresponde a los nt 978-958, aminoácidos 320-326, de Choo *et al.*, PNAS 1991.

Cuando se prepararon las sondas de ADN de ADNmc para uso en el HTA, se biotiniló uno de los cebadores para la PCR anidada. Véase, por ejemplo, la SEQ ID NO: 6 en la Tabla 1.

TABLA 1

HCV-1	nt 5 '→ 3'	nt 3' → 5'	Aminoácido	Tipo de cebador
(SEQ ID NO: 1) Purificado C170S	508-529	529-508	170-176	Sonda con sentido universal, PCR I
(SEQ ID NO: 2) Purificado, E338A1	1032-1012	1012-1032	338-344	Antisentido de tipo I, PCR I
(SEQ ID NO: 3) Purificado, E338A2a	1032-1012	1012-1032	338-344	Antisentido de tipo 2a, PCR I
(SEQ ID NO: 4) Purificado, E338A2b	1032-1012	1012-1032	338-344	Antisentido de tipo 2b, PCR I
(SEQ ID NO: 5) Purificado, E338A3a	1032-1012	1012-1032	338-344	Antisentido de tipo 3a, PCR I
(SEQ ID NO: 6) Purificado, E320A	978-958	958-978	320-326	Antisentido universal, PCR II
(SEQ ID NO: 7) Purificado, C179S1	536-557	958-978	179-186	Con sentido de tipo 1, PCR II
(SEQ ID NO: 8) Purificado, C179S2	536-557	958-978	179-186	Con sentido de tipo 2, PCR II
(SEQ ID NO: 9) Purificado, C179S3	536-557	958-978	179-186	Con sentido de tipo 3, PCR II

Ejemplo 3

HTA

Se prepararon las sondas monocatenarias mediante PCR-TI de sueros positivos por ELISA de HCV y RIBA de genotipos conocidos con los mismos cebadores de PCR descritos como anteriormente, excepto porque uno de los cebadores 320A estaba biotinilado. Se generaron sondas de ADNmc con Dynabeads M-280 Streptavidin siguiendo el protocolo de Heng Pan y Eric Delwart. Se eluyó la cadena simple no biotinilada de la columna de perlas magnéticas/estreptavidina. Se generaron sondas a partir de 20 ng de ADNmc de los diferentes genotipos y se marcaron terminalmente utilizando polinucleótido cinasa T4 (Gibco BRL) y 100 μCi de ^{32}P-ATP, y se purificaron después en columna. Se separó la sonda de cinasa del ^{32}P-ATP utilizando una columna Bio-Sepharose de Pharmacia. Se mezclaron las sondas monocatenarias marcadas con ^{32}P con un exceso de 100 veces de promotor, y se generaron los productos de PCR a partir de las muestras de paciente o suero/plasma de control. La hibridación fue en 2 x SSC. Se pusieron las mezclas en un bloque de calentamiento a 94ºC durante 3 minutos. Se transfirieron después a un bloque de calentamiento a 55ºC durante al menos 2 horas. Se cargó todo el volumen de reacción en un gel MDE de poliacrilamida al 6% de 1 mm de grosor (Baker) y se sometió a electroforesis durante 16 h a 500 V. Se secó a vacío el gel a 80ºC en papel de filtro y se expuso a película de rayos X. Se determinaron los genotipos de cada una de las muestras basándose en el método de Delwart. La Tabla 2 exhibe los resultados de genotipo determinados utilizando el HTA.

Las Figuras 1A-1E son autorradiogramas que muestran cada una de las sondas monocatenarias de la Tabla 1, es decir, las sondas específicas conocidas para los genotipos 1a, 1b, 2b, 3a en las Figuras 1A-E, respectivamente, véase el carril al extremo izquierdo del gel MDE. Se muestra el homodúplex (h) (sonda mc al producto de PCR-TI bicatenario del que se derivó) adyacente a la sonda. Los productos de PCR-TI de los 15 pacientes de diálisis (nº 1, 2, 3, 4, 7, 18, 20, 22, 23, 24, 26, 28, 30, 33, 35) hibridados con la sonda se denominan también en el carril apropiado de cada figura.

Como puede observarse en las Figuras 1A-1E, las sondas de subtipos mc de tipo 1 eran específicas de cada subtipo de tipo 1 y no hibridaban cruzadamente con otros subtipos 1b, 2a, 2b, 3a (2a, 2b no mostrados). La sonda específica de subtipo mc de tipo 3a era también específica del subtipo 3a y no hibridaba cruzadamente con los aislamientos de 1a, 2c o 2a, 2s (datos no mostrados). Las sondas mc de subtipo 2 no hibridan cruzadamente entre sí (datos no mostrados), pero hibridaron cruzadamente con aislamientos de subtipo 2c; sin embargo, la distancia entre el homodúplex y los aislamientos 2c indica un alto grado de divergencia, sugiriendo que los pacientes 23, 30 y 33 tenían diferentes subtipos. Se confirmó mediante secuenciación de la E1 parcial que el virus en los sueros 23, 30 y 33 estaba muy estrechamente relacionado con el subtipo 2c (véase la figura 2b), pero era ambiguo por secuenciación de 5'UTR, véase la Figura 2D.

Los aislamientos 23, 30 y 33 hibridaban con la sonda 2a, aunque sólo 30 y 33 hibridaban con la sonda 2b. Los geles indican también que el aislamiento 30 está más estrechamente relacionado con 2a que con 2b. Por lo tanto, aunque los tres sueros son claramente de subtipo 2 no a no b, no son igualmente divergentes de los tipos 2a y 2b. Como se observa en las Figuras 1B y 1D, el paciente 4 parece estar coinfectado con los tipos 1b y un subtipo de tipo no a no b.

La sonda 1b derivaba de un paciente (JK 16) y parecía tener dos genomas víricos, lo que se refleja en el carril de homodúplex (h), y por lo tanto cada paciente 1b tiene dos bandas.

La sonda 3a mc derivaba de un clon plasmídico de un producto de PCR-TI de un individuo de tipo 3a (JK3a), véase la Fig. 1E, carril h, por lo tanto, las múltiples bandas en el carril 22 reflejan lo más probablemente dos virus estrechamente relacionados en este paciente.

Parecía que la mayoría de las veces los pacientes tenían aislamientos víricos únicos. Es posible que los pacientes 3 y 18 tuvieran aislamientos víricos idénticos o altamente relacionados. De forma similar, los pacientes 20 y 26 tenían el mismo tipo de aislamiento vírico de tipo 3a y los pacientes 2 y 4 tienen el mismo tipo de aislamiento 1b basándose en la comigración de las bandas en geles MDE.

Las Figuras 2a-2c exhiben árboles filogenéticos, dendrogramas, que muestran las relación genética de cada una de las secuencias nucleotídicas de E1 parcial. Estos dendogramas se construyeron mediante alineamiento progresivo pareado de las secuencias nucleotídicas entre sí utilizando el programa de software informático "Gene Works Unweighted Pair Group Methods" con media aritmética, como se describe en Weiner *et al*., J. Virol. 67: pág. 4365-4368 (1993). Los dendogramas en las Figuras 2a-2c se formaron comparando secuencias de E1 parciales de supuestos aislamientos de tipo 1 (nt 625-93), de tipo 2 (nt 583-915) o tipo 3 (nt 558-834) de pacientes de diálisis, como se determina mediante análisis de secuencia de secuencias de genotipo publicadas para el tipo 1a (HCV-1) (Choo *et al.*, PNAS 1991); 1b (HCV-J) (Kato *et al.*); 2a (HC-J6 (Okamoto *et al.* (1992)); 2b (HC-J8) (Okamoto *et al.*, 1992); 2c (Bukh *et al.*, PNAS 1993) y 3a (NZL-1) (Sakamoto *et al.*, 1994) en la misma región del genoma.

La Figura 2D es un dendrograma formado como se describe anteriormente comparando secuencias 5'UTR parciales de los aislamientos 23, 30 y 33 con secuencias de genotipo publicado de tipo 1, 2 y 3 (nt-274 a -81) para la misma región del genoma.

Se compararon los resultados de RFLP y HTA y se presentan en la Tabla 2.

TABLA 2

Comparación de los resultados de genotipado por HTA y RFLP de E1 parcial

Paciente	HTA	RFLP
1	1b	1b
2	1b	1b
3	3a	3a
4	1b	1b
7	3a	3a
18	3a	3a
20	3a	3a
22	3a	3a
23	2?*	2a
24	1b	1b
26	3a	3a
28	1b	1b
30	2?*	2a
33	2?*	2a
35	3a	3a

* la muestra no es 2a ni 2b

Las secuencias de E1 parciales exhibidas en las Figuras 3a-3d confirman las denominaciones de subtipo de HTA dadas en la Tabla 2 y muestran definitivamente que los pacientes 23, 30 y 33 están más estrechamente relacionados con 2c, estando 33 más distantemente relacionado con 2c (18,6% divergente).

Los resultados de RFLP utilizando ScrFI (véase Davidson *et al.*, J. Gen. Virol. (1995) 76: 1197-1204) denominaron erróneamente a 23, 30 y 33 como de tipo 2a. Esta denominación errónea se refleja en la Figura 2D, que muestra que basándose en la secuencia de nt 5'UTR el ordenador no subtipó exactamente HCV 2c debido a una divergencia de nt insuficiente en esta región del genoma.

La presente invención de HTA que utiliza cebadores para la región de núcleo y cubierta permitió 3 niveles de caracterización de genomas de HCV. El primero era la especificidad de tipo en la elección de los cebadores de PCR-TI. El segundo era la especificidad de subtipo, basada en elegir cebadores en la región de núcleo/E1 y de una región mayor de 400 nt, lo que dio como resultado la falta de hibridación cruzada entre sondas de subtipo, por ejemplo, 1 y 3, 2a, 2b; y un alto grado de heterogeneidad para maximizar las diferencias entre genotipos (falta de hibridación cruzada). Finalmente, se determinó la especificidad de aislamiento mediante la distancia desde el homodúplex como se ejemplifica en las Figuras 1A-1E. Otros métodos de genotipado no tienen la capacidad de analizar diferencias de aislamiento.

REIVINDICACIONES

1. Un oligonucleótido consistente en la secuencia de SEQ ID NO: 1.

2. Un oligonucleótido consistente en la secuencia de SEQ ID NO: 2.

3. Un oligonucleótido consistente en la secuencia de SEQ ID NO: 3.

4. Un oligonucleótido consistente en la secuencia de SEQ ID NO: 4.

5. Un oligonucleótido consistente en la secuencia de SEQ ID NO: 5.

6. Un oligonucleótido consistente en la secuencia de SEQ ID NO: 6.

7. Un oligonucleótido consistente en la secuencia de SEQ ID NO: 7.

8. Un oligonucleótido consistente en la secuencia de SEQ ID NO: 8.

9. Un oligonucleótido consistente en la secuencia de SEQ ID NO: 9.

10. Un par de cebadores de PCR en los que el cebador con sentido consiste en la SEQ ID NO: 1 y el cebador antisentido se selecciona del grupo consistente en las SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4 y SEQ ID NO: 5.

11. Un par de cebadores de PCR en los que el cebador antisentido consiste en la SEQ ID NO: 6 y el cebador con sentido se selecciona del grupo consistente en las SEQ ID NO: 7, SEQ ID NO: 8 y SEQ ID NO: 9.

12. Un método de determinación del genotipo del HCV de una cepa de HCV, comprendiendo dicho método las etapas de:

(a) someter dicha cepa de HCV a una o más etapas de PCR, en el que las una o más etapas de PCR utilizan una sonda con sentido del núcleo o de la región E1 del genoma del HCV y una sonda antisentido del núcleo o de la región E1 del genoma del HCV;

(b) formar un heterodúplex mediante la desnaturalización y reasociación de mezclas del producto amplificado obtenido en la etapa (a) con fragmentos de ADN o ARN de un genotipo de HCV conocido;

(c) comparar la movilidad de dicho heterodúplex en un sistema que separa por tamaño con la movilidad de un homodúplex de fragmentos de ADN o ARN de genotipo conocido para determinar el genotipo de la cepa de HCV.

13. El método de la reivindicación 12, en el que dicha cepa de HCV se somete a dos etapas de PCR, en el que el primer conjunto de cebadores comprende una sonda con sentido universal de las regiones de núcleo o E1 del genoma de HCV y una sonda antisentido específica de tipo de las regiones de núcleo o E1 del genoma de HCV, y en el que el segundo conjunto de cebadores de PCR comprende una sonda antisentido universal de las regiones de núcleo o E1 del genoma de HCV y una sonda con sentido específica de tipo de las regiones de núcleo o E1 del genoma de HCV.

14. El método de la reivindicación 12, en el que el primer conjunto de cebadores de PCR son aquellos según la reivindicación 10 y en el que el segundo conjunto de cebadores de PCR son aquellos según la reivindicación 11.

15. El método de la reivindicación 12, en el que dichas fragmentos de ADN o ARN de un genotipo conocido comprenden una sonda de ADN.

16. El método de la reivindicación 15, en el que dicha sonda es monocatenaria.

17. El método de la reivindicación 16, en el que dicha sonda de ADN está radiomarcada.

18. El método de la reivindicación 16, en el que dicha sonda de ADN monocatenaria se obtiene mediante amplificación por PCR.

19. El método de la reivindicación 18, en el que dicha sonda de ADN se obtiene mediante amplificación de PCR en dos etapas utilizando los cebadores de la reivindicación 10 para la primera etapa y de la reivindicación 11 para la segunda etapa.

20. El método de la reivindicación 12, en el que dicha cepa de HCV está presente en exceso en la mezcla que forma el heterodúplex.

21. Un método para predecir la respuesta a la terapia farmacológica de una cepa de HCV de un paciente infectado con dicha cepa de HCV, comprendiendo dicho método determinar la sensibilidad de genotipos de HCV conocidos

frente a dicha terapia farmacológica, determinar el genotipo de HCV de dicha cepa de HCV mediante el método según la reivindicación 12, y comparar dicho genotipo de HCV de dicha cepa antes de dicha terapia farmacológica con dicha sensibilidad de genotipos de HCV conocidos frente a dicha terapia farmacológica.

22. Un método para predecir la respuesta a una vacuna terapéutica de una cepa de HCV de un paciente infectado con dicha cepa de HCV, comprendiendo dicho método determinar la sensibilidad de genotipos de HCV conocidos frente a dicha vacuna terapéutica, determinar el genotipo de HCV de dicha cepa de HCV mediante el método según la reivindicación 12, y comparar dicho genotipo de HCV de dicha cepa antes de la administración de dicha vacuna terapéutica con dicha sensibilidad de genotipos de HCV conocidos frente a dicha vacuna terapéutica.

23. Un método para predecir la idoneidad de una composición de vacuna profiláctica para una población de muestra dada, comprendiendo dicho método determinar el genotipo de dicha vacuna profiláctica, determinar la predominancia de genotipos de HCV conocidos en dicha población de muestra mediante el método según la reivindicación 12, y comparar dicho genotipo de HCV de dicha cepa de vacuna profiláctica con el genotipo predominante determinado antes de la administración de dicha vacuna profiláctica a dicha muestra de población.

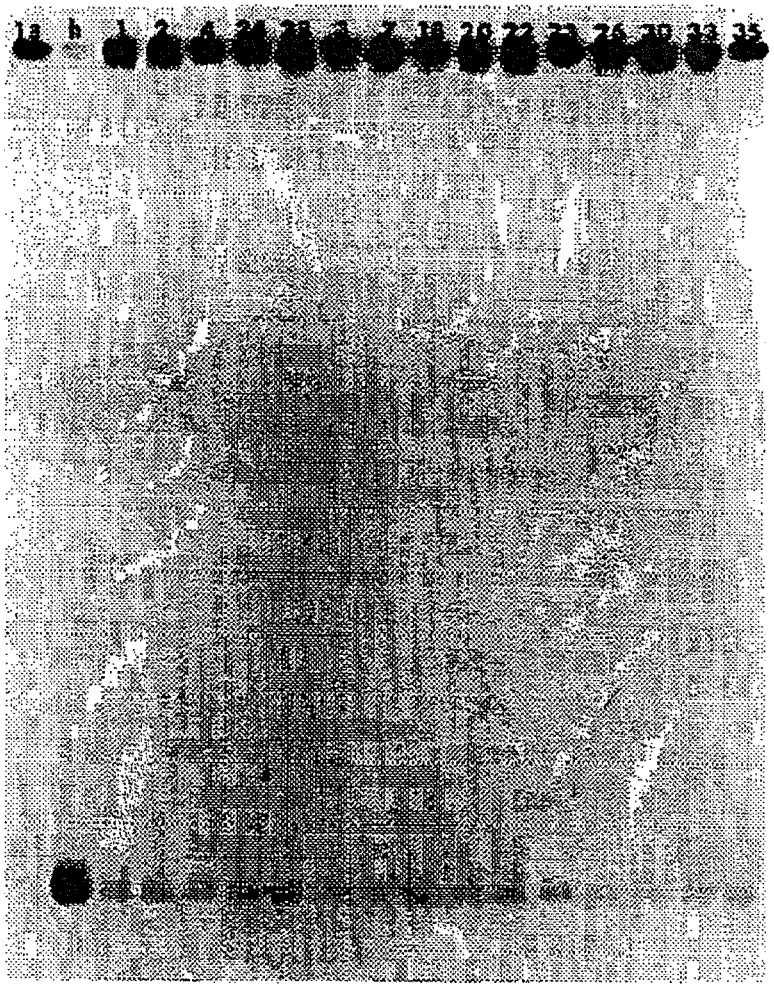

FIG. 1A

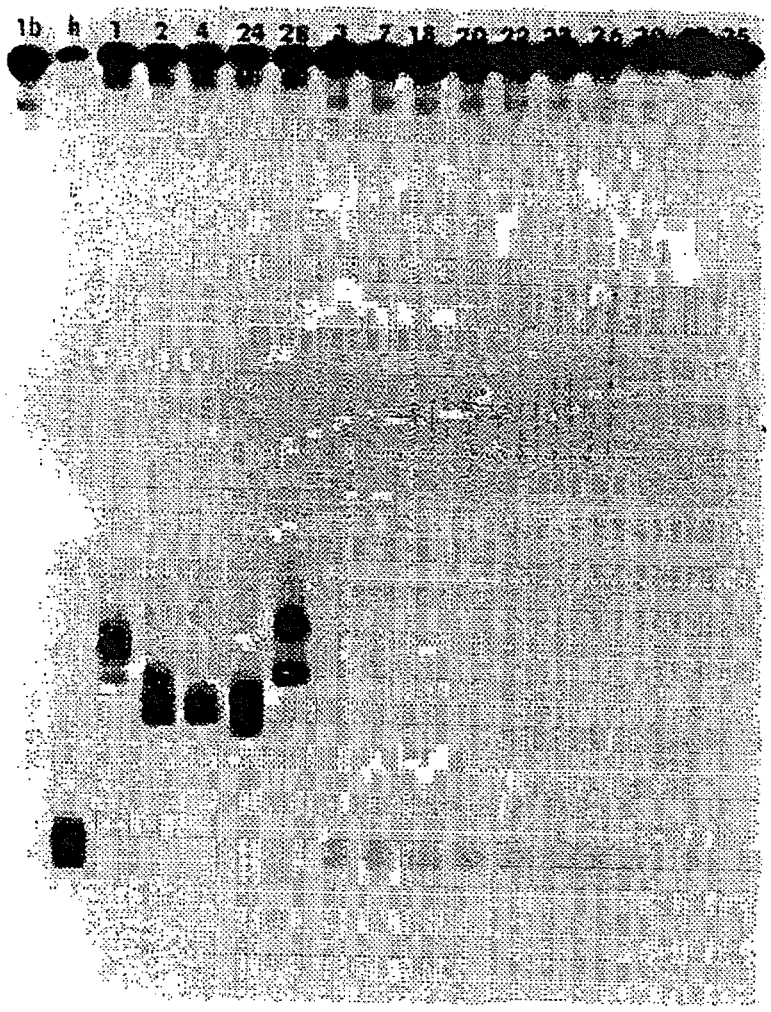

FIG. 1B

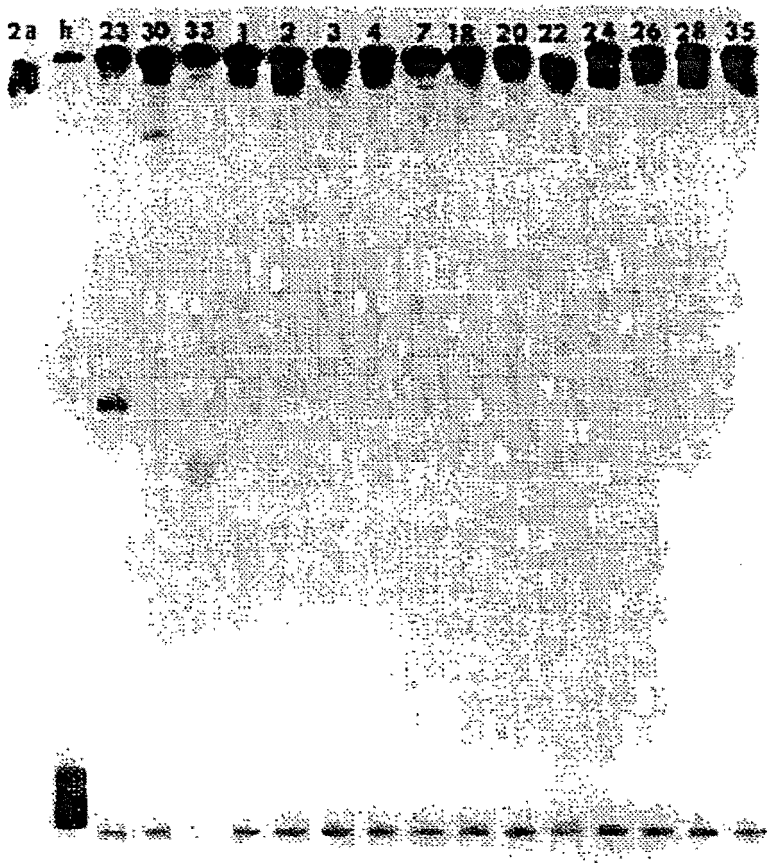

FIG. 1C

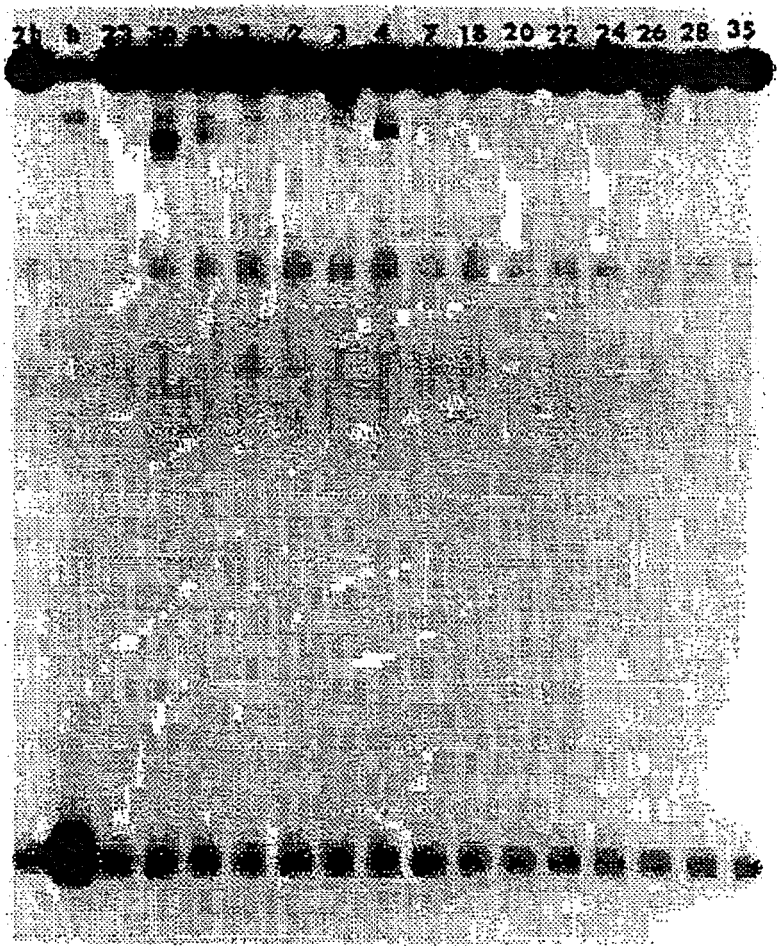

FIG. 1D

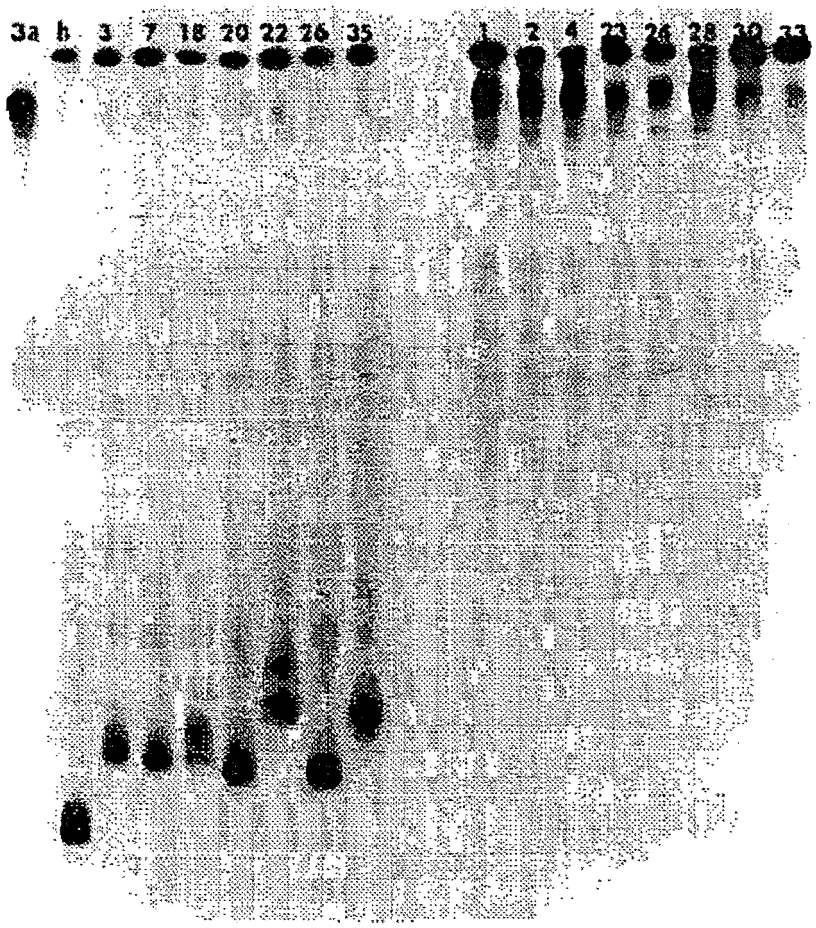

FIG. 1E

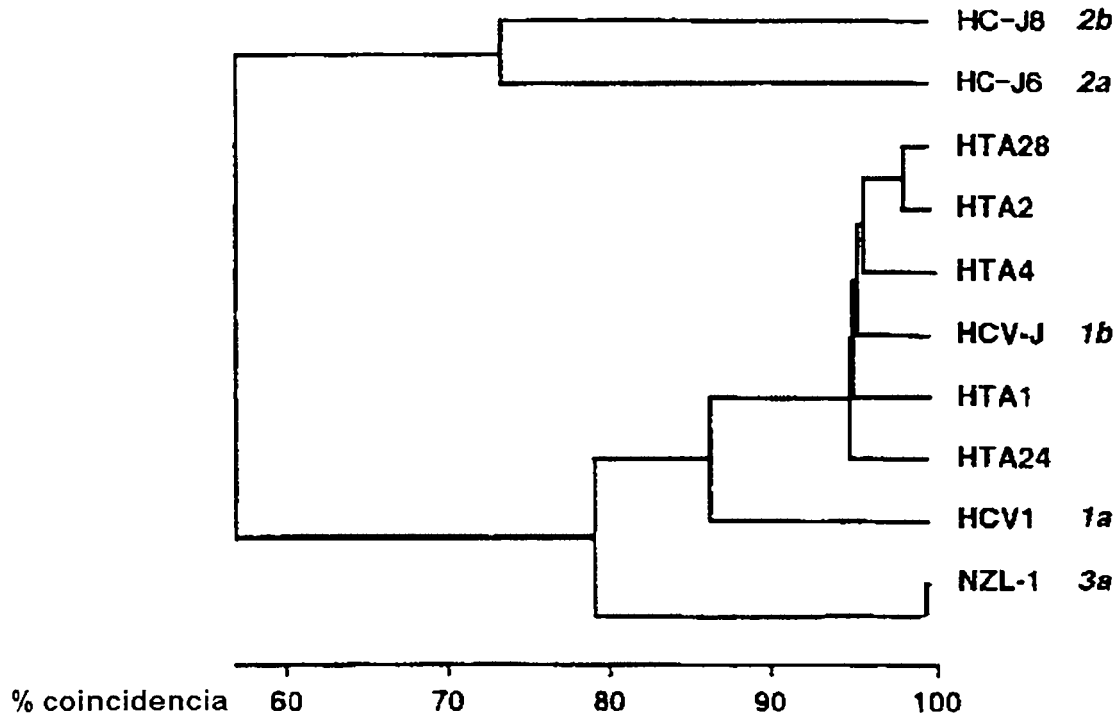
HC-J8 2b
HC-J6 2a
HTA28
HTA2
HTA4
HCV-J 1b
HTA1
HTA24
HCV1 1a
NZL-1 3a
% coincidencia
60
70
80
90
100

FIG. 2A

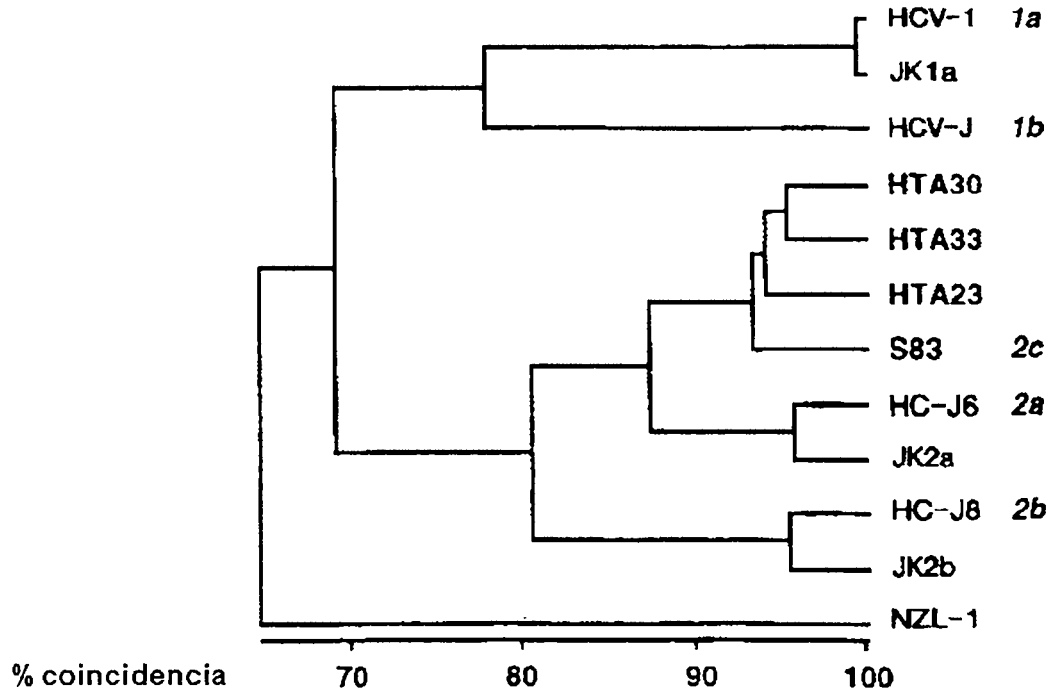
HCV-1 1a
JK1a
HCV-J 1b
HTA30
HTA33
HTA23
S83 2c
HC-J6 2a
JK2a
HC-J8 2b
JK2b
NZL-1
% coincidencia
70
80
90
100

FIG. 2B

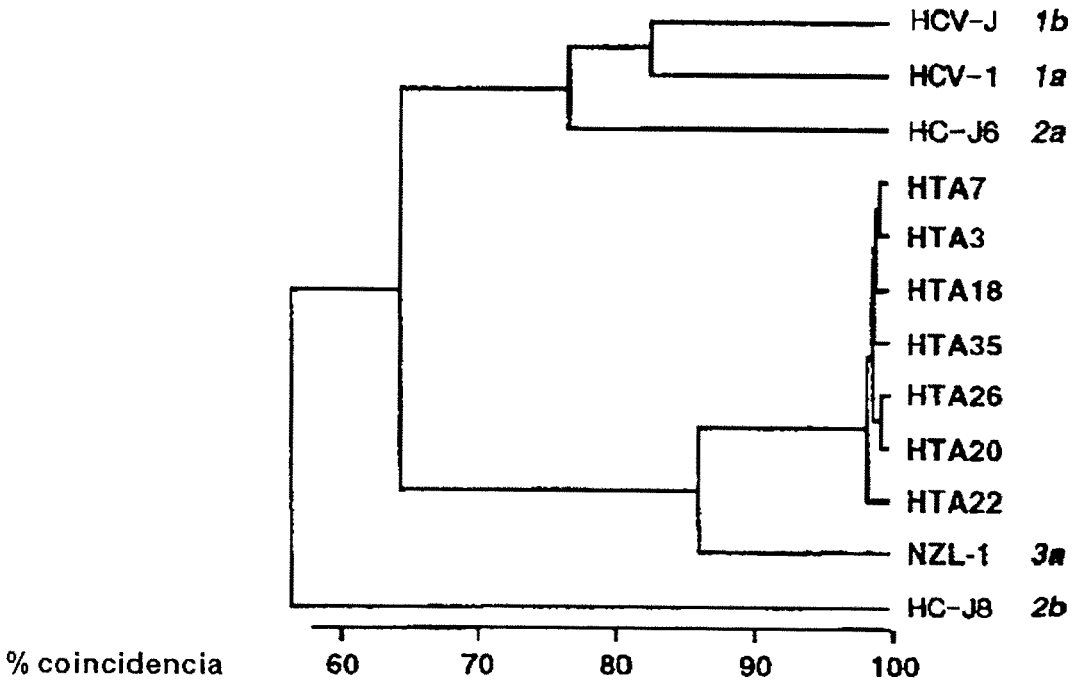
HCV-J 1b
HCV-1 1a
HC-J6 2a
HTA7
HTA3
HTA18
HTA35
HTA26
HTA20
HTA22
NZL-1 3a
HC-J8 2b
% coincidencia
60
70
80
90
100

FIG. 2C

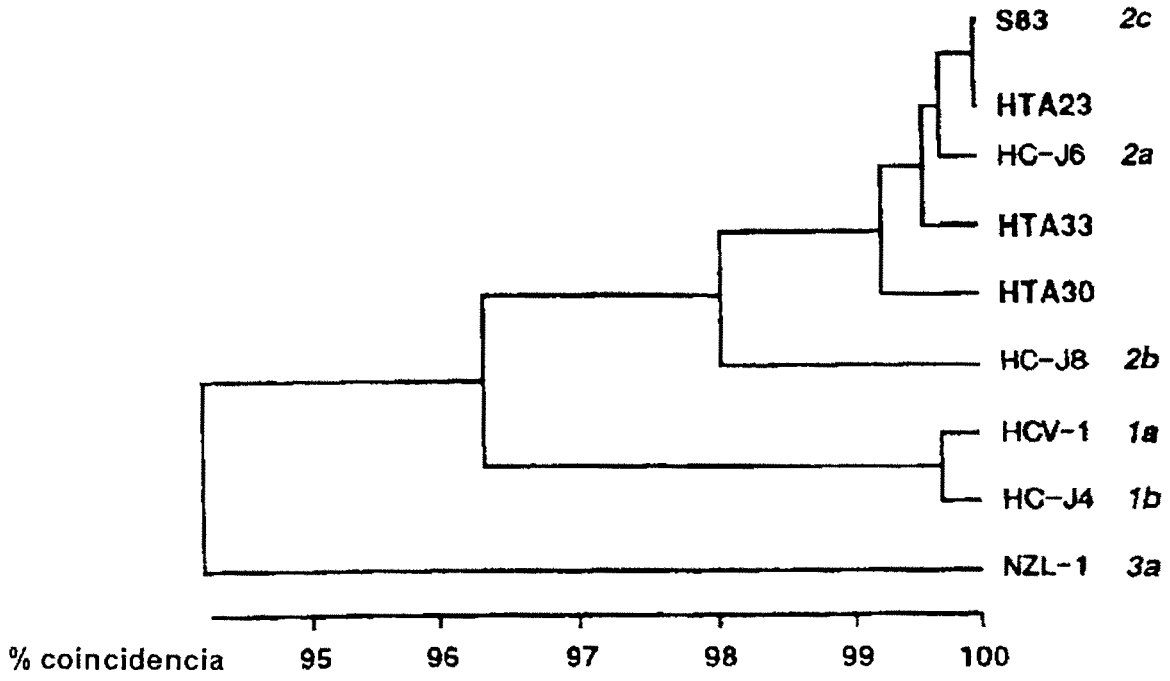
S83 2c
HTA23
HC-J6 2a
HTA33
HTA30
HC-J8 2b
HCV-1 1a
HC-J4 1b
NZL-1 3a
% coincidencia
95
96
97
98
99
100

FIG. 2D

```
HTA1      AACTCAAGCATTGTGTATGAAGCGGCGGACATGATCATGCACACCCCCGGGTGCGTGCCA
HTA2      AACTCAAGCATCGTGTATGAGGCAGCGGAAGTGATCATGCACATTCCCGGGTGCGTGCCC
HTA4      AACTCAAGCATAGTATATGAGGCAGCGGACATAATCATGCATACCCCCGGGTGCGTGCCC
HTA24     AACACGAGCATTGTGTATGAGGCAGCGGACTTGATCATGCACGTCCCCGGGTGCGTGCCC
HTA28     AACTCAAGCATCGTGTATGAGGCGGCGGAAGTGATCATGCACATTCCTGGGTGCGTGCCC
HCV-1     AACTCGAGTATTGTGTACGAGGCGGCCGATGCCATCCTGCACACTCCGGGGTGCGTCCCT
HCV-J     AACTCAAGTATTGTGTATGAGGCAGCGGACATGATCATGCACACCCCCGGGTGCGTGCCC
HC-J6     AATGATAGCATTACCTGGCAACTCCAGGCTGCTGTCCTCCACGTCCCCGGGTGCGTCCCG
HC-J8     AACAACAGCATCACCTGGCAGCTCACTGACGCAGTTCTCCATCTTCCTGGATGCGTCCCA
NZL-1     AATAGCAGTATTGTGTATGAGGCCGATGATGTCATTCTGCACACACCCGGCTGTGTACCT

HTA1      TGCGTCCGGGAGGGCAATCTCTCCCGCTGCTGGGTAGCGCTCACTCCCACGCTCGCGGCC
HTA2      TGCGTTCGGGAGAGCAATCTCTCCCGCTGCTGGGTAGCGCTCACCCCCACACTCGCGGCC
HTA4      TGTGTTCGGGAGGTCAACTCCTCCCGCTGCTGGGCAGCGCTCACCCCTACGCTCGCGGCC
HTA24     TGCGTTCGGGAGGGCAACTCCTCCCGATGCTGGGTAGCGCTCACTCCCACGATCGCGGCC
HTA28     TGCGTTCGGGAGGGCGACTTCTCCCGCTGCTGGGTAGCGCTCACCCCCACACTCGCGGCC
HCV-1     TGCGTTCGTGAGGGCAACGCCTCGAGGTGTTGGGTGGCGATGACCCCTACGGTGGCCACC
HCV-J     TGCGTCCGGGAGAGTAATTTCTCCCGTTGCTGGGTAGCGCTCACTCCCACGCTCGCGGCC
HC-J6     TGCGAGAAAGTGGGGAATACATCTCGGTGCTGGATACCGGTCTCACCGAATGTGGCCGTG
HC-J8     TGTGAGAATGATAATGGCACCTTGCATTGCTGGATACAAGTAACACCCAACGTGGCTGTG
NZL-1     TGTGTCCAGGACGGCAATACATCTACGTGCTGGACCCCAGTGACACCTACAGTGGCAGTC
JK1a      TGTGTTCGCGAGGGCAACGCCTCGAGGTGTTGGGTGGCGATGACCCCTACGGTGGCCACC

HTA1      AGAAACAGCAGCGTTCCTACTACGACAATACGACGCCATGTCGACTTGCTAGTAGGAGCG
HTA2      AGGAACAGCAGCGTCCCCACCACGACAATACGACGCCACGTCGACTTGCTCGTTGGGGCG
HTA4      AGGAACTCCAGCGTGCCCACTACGACAATACGACGCCACGTCGACTTGCTCGTTGGGGCG
HTA24     AGGAACAGCAGTGTCCCCGTTACGACCATACGACGCCACGTCGATTTGCTCGTTGGGGCG
HTA28     AGGAATAACAGCGTCCCCACTACGACAATACGACGCCACGTCGACTTGCTCGTTGGGGCG
HCV-1     AGGGATGGCAAACTCCCCGCGACGCAGCTTCGACGTCACATCGATCTGCTTGTCGGGAGC
HCV-J     AGGAACAGCAGCATCCCCACCACGACAATACGACGCCACGTCGATTTGCTCGTTGGGGCG
HC-J6     CAGCAGCCCGGCGCCCTCACGCAGGGCTTACGGACGCACATTGACATGGTTGTGATGTCC
HC-J8     AAACACCGCGGTGCGCTCACTCGTAGCCTGCGAACACACGTCGACATGATCGTAATGGCA
NZL-1     AGGTACGTCGGAGCAACTACTGCTTCGATACGCAGTCATGTGGACCTATTAGTAGGCGCG

HTA1      GCTGCTTTTTGCTCCGCCATGTACGTGGGGGACCTCTGCGGATCTATTTTCCTCGTCTCC
HTA2      GCTGCCTTCTGCTCCGCTATGTATGTGGGGGATCTCTGCGGATCTGTTTTCCTTGTCTCC
HTA4      GCTGCTTTCTGCTCCGCTATGTACGTGGGGGATCTATGCGGATCTGTTCTACTTGTCTCT
HTA24     GCTGCTCTTTGCTCCGCCATGTACGTGGGGGATCTCTGCGGATCTGTCTTCCTCGCTTCC
HTA28     GCTGCCTTCTGCTCCGCTATGTACGTGGGGGATCTCTGCGGATCTGTTTTCCTTGTCTCC
HCV-1     GCCACCCTCTGTTCGGCCCTCTACGTGGGGGACCTATGCGGGTCTGTCTTTCTTGTCGGC
HCV-J     GCTGCTCTCTGTTCCGCTATGTACGTTGGGGATCTCTGCGGATCCGTTTTTCTCGTCTCC
HC-J6     GCCACGCTCTGCTCCGCTCTTTACGTGGGGGACCTCTGCGGTGGGGTGATGCTTGCAGCC
HC-J8     GCTACGGCCTGCTCGGCCTTGTATGTGGGAGATGTGTGCGGGGCCGTGATGATTCTATCG
NZL-1     GCCACGATGTGCTCTGCGCTCTACGTGGGTGATATGTGTGGGGCTGTCTTTCTCGTGGGA

HTA1      CAACTGTTCACCTTCTCGCCCCGCCGGCATCATACAGTACAGGACTGCAATTGCTCGATC
HTA2      CAACTGTTCACCTTTTCGCCTCGCCGGCATGAGACAGTACAGGACTGCAATTGTTCAATC
HTA4      CAGCTGTTCACCTTCTCACCTCGCCGGCACGAGACAGTGCAGGACTGCAATTGTTCAATC
HTA24     CAGTTGTTCACTTTCTCGCCTCGCCAGCATCAGACGGTACAGGACTGCAACTGCTCAATC
HTA28     CAACTGTTCACCTTTTCGCCTCGCCGGCATGCGACAGTACAGGACTGCAATTGTTCAATC
```

FIG. 3A-1

```
HCV-1       CAACTGTTCACCTTCTCTCCCAGGCGCCACTGGACGACGCAAGGTTGCAATTGCTCTATC
HCV-J       CAGCTGTTCACCTTCTCACCTCGCCGGTATGAGACGGTACAAGATTGCAATTGCTCAATC
HC-J6       CAGATGTTCATTGTCTCGCCACAGCACCACTGGTTTGTGCAAGACTGCAATTGCTCCATC
HC-J8       CAGGCTTTCATGGTATCACCACAACGCCACAACTTCACCCAAGAGTGCAACTGTTCCATC
NZL-1       CAAGCCTTCACGTTCAGACCTCGACGCCATCAAACGGTCCAGACCTGTAACTGCTCGCTG

HTA1        TATCCC
HTA2        TATCCC
HTA4        TATCCC
HTA24       TATCCC
HTA28       TATCCC
HCV-1       TATCCC
HCV-J       TATCCC
HC-J6       TACCCT
HC-J8       TACCAA
NZL-1       TACCCA
```

FIG. 3A-2

```
JK1a        CGCAACTCCACGGGGCTTTACCACGTCACCAATGATTGCCCTAACTCGAGTATTGTGTAC
JK2a        AAGAACACCAGCGACAGCTACATGGTGACCAATGACTGCCAAAATGACAGCATCACCTGG
JK2b        AGGAACATCAGTTCTAGCTACTACGCCACTAATGACTGCTCGAACAACAGCATCACCTGG
SW83.2c     AAGGACACCGGCGACTCCTACATGCCGACCAACGATTGCTCCAACTCTAGTATCGTTTGG
HTA23       AAAAACACCAGCATCTCCTATATGGCGACCAACGACTGCTCCAATTCCAGCATCGCTTGG
HTA33       AAGAACACCAGCGACTCCTACATGGCGACTAACGACTGCTCTAACTCCAGCATCGTTTGG
HTA30       AAGAACACCAGCACCTCCTACATGGTGACTAACGATTGCTCCAACTCCAGCATCGTTTGG
HCV-1       CGCAACTCCACGGGGCTTTACCACGTCACCAATGATTGCCCTAACTCGAGTATTGTGTAC
HCV-J       CGCAACGTGTCCGGGATATACCATGTCACGAACGACTGCTCCAACTCAAGTATTGTGTAT
HC-J8       AGGAACATTAGTTCTAGCTACTACGCCACTAATGATTGCTCAAACAACAGCATCACCTGG
NZL-1       CGGAATACGTCTGGCCTCTACGTCCTTACCAACGACTGTTCCAATAGCAGTATTGTGTAT
HC-J6       AAGAACATCAGTACCGGCTACATGGTGACCAACGACTGCACCAATGATAGCATTACCTGG
                 *              **       ** ** ** **     **     ** **     *

JK1a        GAGACGGCCGATGCCATCCTGCACACTCCGGGGTGCGTCCCTTGTGTTCGCGAGGGCAAC
JK2a        CAGCTTGAGGCTGCGGTCCTCCACGTCCCCGGGTGCGTCCCGTGCGAGAGAGTGGGAAAT
JK2b        CAGCTCACCAACGCAGTTCTCCACCTTCCCGGATGCGTCCCATGTGAGAATAATAATGGC
SW83.2c     CAGCTTGAAGGAGCAGTGCTTCATACTCCTGGATGCGTCCCTTGTGAGCGTACCGCCAAC
HTA23       CAGTTTGACGGCGCAGTGCTCCATACTCCTGGATGTGTCCCTTGCGAACGGACCGGCAAC
HTA33       CAGCTTGAGGACGCAGTGCTCCATGTCCCTGGATGTGTCCCTTGTGAGAAGACTGGCAAT
HTA30       CAACTTGAAGGCGCAGTGCTCCATGTTCCTGGATGTGTCCCTTGTGAGCAGATCGGCAAC
HCV-1       GAGGCGGCCGATGCCATCCTGCACACTCCGGGGTGCGTCCCTTGCGTTCGTGAGGGCAAC
HCV-J       GAGGCAGCGGACATGATCATGCACACCCCCGGGTGCGTGCCCTGCGTCCGGGAGAGTAAT
HC-J8       CAGCTCACTGACGCAGTTCTCCATCTTCCTGGATGCGTCCCATGTGAGAATGATAATGGC
NZL-1       GAGGCCGATGATGTCATTCTGCACACACCCGGCTGTGTACCTTGTGTCCAGGACGGCAAT
HC-J6       CAACTCCAGGCTGCTGTCCTCCACGTCCCCGGGTGCGTCCCGTGCGAGAAAGTGGGGAAT
             *                 *  * **    ** ** ** ** ** ** *

JK1a        GCCTCGAGGTGTTGGGTGGCGATGACCCCTACGGTGGCCACCAGGGATGGCAAACTCCCC
JK2a        ACATCTCGGTGCTGGATACCGGTCTCACCAAACGTGGCTGTGCGGCAGCCCGGCGCCCTC
JK2b        ACCTTGCATTGCTGGATACAAGTAACACCTAATGTGGCCGTAAAACATCGCGGCGCACTC
SW83.2c     GTCTCTCGATGTTGGGTGCCGGTTGCCCCCAATCTCGCCATAAGTCAACCTGGCGCTCTC
HTA23       GCGTCCCGGTGTTGGGTGCCGGTTGCCCCCAATGTGGCTATAAGACAACCCGGCGCCCTC
HTA33       ACGTCTCGGTGCTGGGTGCCGGTTACCCCCAATGTGGCTACAAGTCAACCCGGCGCTCTC
HTA30       GTGTCTCAGTGTTGGGTGCCGGTTACCCCCAATATGGCCATAAGTACACCCGGCGCTCTC
HCV-1       GCCTCGAGGTGTTGGGTGGCGATGACCCCTACGGTGGCCACCAGGGATGGCAAACTCCCC
HCV-J       TTCTCCCGTTGCTGGGTAGCGCTCACTCCCACGCTCGCGGCCAGGAACAGCAGCATCCCC
HC-J8       ACCTTGCATTGCTGGATACAAGTAACACCCAACGTGGCTGTGAAACACCGCGGTGCGCTC
NZL-1       ACATCTACGTGCTGGACCCCAGTGACACCTACAGTGGCAGTCAGGTACGTCGGAGCAACT
HC-J6       ACATCTCGGTGCTGGATACCGGTCTCACCGAATGTGGCCGTGCAGCAGCCCGGCGCCCTC
               *     ** ***        *  * ** *   * **

JK1a        GCGACGCAGCTTCGACGTCACATCGATCTGCTTGTCGGGAGCGCCACCCTCTGTTCGGCC
JK2a        ACGCAGGGCTTGCGGACGCACATCGACATGATTGTGATGTCCGCCACGCTCTGCTCCGCT
JK2b        ACTCACAACCTGCGGACACATGTCGACATGATCGTAATGGCAGCTACGGTCTGTTCGGCC
SW83.2c     ACTAAGGGCCTGCGAGCACACATCGATATCATCGTGATGTCTGCTACGGTCTGTTCTGCC
HTA23       ACTAAGGGCATACGAACGCACATTGATGTCATCGTAATGTCTGCTACGCTCTGTTCTGCC
HTA33       ACCAGGGGCTTGCGGACGCACATCGATGTCATCGTGATGTCAGCCACGCTCTGCTCCGCT
HTA30       ACTAAGGGCTTGCGAACGCACATCGACGGCATCGTGATGTCCGCTACGCTCTGTTCTGCC
HCV-1       GCGACGCAGCTTCGACGTCACATCGATCTGCTTGTCGGGAGCGCCACCCTCTGTTCGGCC
HCV-J       ACCACGACAATACGACGCCACGTCGATTTGCTCGTTGGGGCGGCTGCTCTCTGTTCCGCT
HC-J8       ACTCGTAGCCTGCGAACACACGTCGACATGATCGTAATGGCAGCTACGGCCTGCTCGGCC
NZL-1       ACTGCTTCGATACGCAGTCATGTGGACCTATTAGTAGGCGCGGCCACGATGTGCTCTGCG
HC-J6       ACGCAGGGCTTACGGACGCACATTGACATGGTTGTGATGTCCGCCACGCTCTGCTCCGCT
             *        * **    **  * **     * **        **   *    ** ** **

JK1a        CTCTACGTGGGGGATCTGTGCGGGTCTGTCTTTCTTGTCGGCCAACTGTTTACCTTCTCT
```

FIG. 3B-1

```
JK2a        CTCTACGTGGGGGACCTCTGTGGCGGGATGATGCTCGCAGCCCAGATGTTCATCGTTTCG
JK2b        TTGTACGTAGGAGACGTGTGTGGGGCTGTGATGATTGTGTCTCAGGCCCTTATAATATCA
SW83.2c     CTTTATGTGGGGGACGTGTGTGGCGCGCTGATGCTGGCCGCTCAGGTCGTCGTCGTGTCG
HTA23       CTTTACGTGGGGGACGTGTGTGGTGCGCTGATGATTGCCGCTCAGGTCGTCATTGTGTCT
HTA33       CTCTATGTGGGGGACGTGTGTGGCGCGTTGACGATAGCCGCTCAGGTTGTCATCGTATCG
HTA30       CTTTATGTGGGGGACGTGTGTGGCGCGTTGATGATAGCCGCCCAGGTCGTCATCGTATCG
HCV-1       CTCTACGTGGGGGACCTATGCGGGTCTGTCTTTCTTGTCGGCCAACTGTTCACCTTCTCT
HCV-J       ATGTACGTTGGGGATCTCTGCGGATCCGTTTTTCTCGTCTCCCAGCTGTTCACCTTCTCA
HC-J8       TTGTATGTGGGAGATGTGTGCGGGGCCGTGATGATTCTATCGCAGGCTTTCATGGTATCA
NZL-1       CTCTACGTGGGTGATATGTGTGGGGCTGTCTTTCTCGTGGGACAAGCCTTCACGTTCAGA
HC-J6       CTTTACGTGGGGGACCTCTGCGGTGGGGTGATGCTTGCAGCCCAGATGTTCATTGTCTCG
             * ** ** ** **  * ** **     *     *       **     *     *

JK1a        CCCAGGCGCCACTGGACGACGCAAGGTTGCAAT
JK2a        CCGCAGAACCACTGGTTCGTGCAGGAATGCAAT
JK2b        CCAGAACACCATAACTTCACCCAAGAGTGCAAC
SW83.2c     CCACAACACCATACGTTTGTCCAGGAATGCAAC
HTA23       CCGCAGCATCACCACTTTGTCCAGGACTGCAAT
HTA33       CCACGGCACCACCACTTTGTCCAGGACTGCAAT
HTA30       CCACAGCACCACCACTTTGTCCACGACTGCAAC
HCV-1       CCCAGGCGCCACTGGACGACGCAAGGTTGCAAT
HCV-J       CCTCGCCGGTATGAGACGGTACAAGATTGCAAT
HC-J8       CCACAACGCCACAACTTCACCCAAGAGTGCAAC
NZL-1       CCTCGACGCCATCAAACGGTCCAGACCTGTAAC
HC-J6       CCACAGCACCACTGGTTTGTGCAAGACTGCAAT
            **       *        **    ** **
```

FIG. 3B-2

```
HTA3        TCATCCAACATCTAGTCTAGAGTGGCGGAATACGTCTGGCCTCTATGTCCTTACCAACGA
HTA7        TCATCCAACATCTAGTCTAGAGTGGCGGAATACGTCTGGCCTCTATGTCCTTACCAACGA
HTA18       TCATCCAACATCTAGTCTAGAGTGGCGGAATACGTCTGGCCTCTATGTCCTTACCAACGA
HTA20       TCATCCAACATCTAGTCTAGAGTGGCGGAATACGTCTGGCCTCTATGTCCTTACCAACGA
HTA22       TCATCCAACATCTAGTCTAGAGTGGCGGAATACGTCTGGCCTCTATGTCCTTACCAACGA
HTA26       TCATCCAACATCTAGTCTAGAGTGGCGGAATACGTCTGGCCTCTATGTCCTTACCAACGA
HTA35       TCATCCAACATCTAGTCTAGAGTGGCGGAATACGTCTGGCCTCTATGTCCTCACCAACGA
NZL-1       TCATCCAGCAGCCAGTCTAGAGTGGCGGAATACGTCTGGCCTCTACGTCCTTACCAACGA
HCV-1       TGTGCCCGCTTCGGCCTACCAAGTGCGCAACTCCACGGGGCTTTACCACGTCACCAATGA
HCV-J       CATCCCAGCTTCCGCTTACGAGGTGCGCAACGTGTCCGGGATATACCATGTCACGAACGA
HC-J6       CACCCCGGTCTCCGCTGCCGAAGTGAAGAACATCAGTACCGGCTACATGGTGACCAACGA
HC-J8       AGTGCCAGTGTCTGCAGTGGAAGTCAGGAACATTAGTTCTAGCTACTACGCCACTAATGA
                 **     *       *       **                **       ** ** **

HTA3        CTGTTCCAATAACATTATTGTGTATGAGGCCGATGACGTCATCCTGCACACGCCCGGCTG
HTA7        CTGTTCCAATAACATCATTGTGTATGAGGCCGATGACGTCATCCTGCACGCACCCGGCTG
HTA18       CTGTTCCAATAATATTATTGTGTATGAGGCCGACGACGTCATCCTGCACGCCCCCGGCTG
HTA20       CTGTTCCAATAACATTATTGTGTATGAGGCCGATGACGTCATCCTGCACACACCCGGCTG
HTA22       CTGTTCCAATAACAGTATTGTGTATGAGGCCGATCACGTCATCCTGCACACACCCGGCTG
HTA26       CTGTTCTAATAACATTATTGTGTATGAGGCCGATGACGTCATCCTGCACACACCCGGCTG
HTA35       CTGTTCCAACAACATTATTGTGTATGAGGCCGATGACGTCATTCTGCACACGCCCGGCTG
NZL-1       CTGTTCCAATAGCAGTATTGTGTATGAGGCCGATGATGTCATTCTGCACACACCCGGCTG
HCV-1       TTGCCCTAACTCGAGTATTGTGTACGAGGCGGCCGATGCCATCCTGCACACTCCGGGGTG
HCV-J       CTGCTCCAACTCAAGTATTGTGTATGAGGCAGCGGACATGATCATGCACACCCCCGGGTG
HC-J6       CTGCACCAATGATAGCATTACCTGGCAACTCCAGGCTGCTGTCCTCCACGTCCCCGGGTG
HC-J8       TTGCTCAAACAACAGCATCACCTGGCAGCTCACTGACGCAGTTCTCCATCTTCCTGGATG
             **  * **    *  **    *   *                *  * **    ** ** **

HTA3        TGTACCTTGTGTTCAGGACGGTAATACATCCAAGTGCTGGACCCCAGTGACACCTACAGT
HTA7        TGTACCTTGTGTTCAGGACGGCAATACATCCACGTGCTGGACCCCAGTGACACCTACAGT
HTA18       TGTACCTTGTGTTCAGGACGGCAATACATCCACGTGCTGGATCCCAGTGACACCTACAGT
HTA20       TGTACCTTGTGTTCAGGACGGCAATACATCCACGTGCTGGACCCCAGTGACACCTACAGT
HTA22       TGTACCTTGTGTTCAAGCCAACAATAAATCCAAATGCTGGACCCCAGTGACACCTACAGT
HTA26       TGTACCTTGTGTTCAGGACGGCAATGCATCCACGTGCTGGACCCCAGTAACACCTACAGT
HTA35       CGTACCTTGTGTACAGGACGGCAATACATCCACGTGCTGGACCCCAGTGACACCTACAGT
NZL-1       TGTACCTTGTGTCCAGGACGGCAATACATCTACGTGCTGGACCCCAGTGACACCTACAGT
HCV-1       CGTCCCTTGCGTTCGTGAGGGCAACGCCTCGAGGTGTTGGGTGGCGATGACCCCTACGGT
HCV-J       CGTGCCCTGCGTCCGGGAGAGTAATTTCTCCCGTTGCTGGGTAGCGCTCACTCCCACGCT
HC-J6       CGTCCCGTGCGAGAAAGTGGGGAATACATCTCGGTGCTGGATACCGGTCTCACCGAATGT
HC-J8       CGTCCCATGTGAGAATGATAATGGCACCTTGCATTGCTGGATACAAGTAACACCCAACGT
             ** ** ** *     *           *     ** ***       *  * ** *   *

HTA3        GGCAGTCAGGTACGTCGGAGCAACCACCGCTTCAATACGCAGCCACGTGGACCTATTATT
HTA7        GGCAGTCAGGTACGTCGGAGCAACCACCGCTTCAATACGCAGCCATGTGGACCTATTAGT
HTA18       GGCAGTCAGGTACGCCGGAGCAACCACCGCTTCAATACGCAGCCATGTGGACCTGTTAGT
HTA20       ATCAGTCAGGTACGTCGGAGCAACCACCGCTTCAATACGCAGCCATGTGGACCTACTATT
HTA22       ATCAGTCGAGTACGTCGGAGCAACCACCGCTTCAATACGCAGCCATGTGGACCTACTATT
HTA26       ATCAGTCAGGTACGTCGGAGCAACCACCGCTTCAGTACGCAGCCATGTGGACCTACTATT
HTA35       GGCAGTCAGGTACGTCGGAGCAACTACCGCTTCAATACGCAGCCATGTGGACCTATTATT
NZL-1       GGCAGTCAGGTACGTCGGAGCAACTACTGCTTCGATACGCAGTCATGTGGACCTATTAGT
HCV-1       GGCCACCAGGGATGGCAAACTCCCCGCGACGCAGCTTCGACGTCACATCGATCTGCTTGT
HCV-J       CGCGGCCAGGAACAGCAGCATCCCCACCACGACAATACGACGCCACGTCGATTTGCTCGT
HC-J6       GGCCGTGCAGCAGCCCGGCGCCCTCACGCAGGGCTTACGGACGCACATTGACATGGTTGT
HC-J8       GGCTGTGAAACACCGCGGTGCGCTCACTCGTAGCCTGCGAACACACGTCGACATGATCGT
              *       *   *         *       * **    **  * **  *  *  *

HTA3        GGGCGCGGCCACGATGTGCTCTGCGCTCTACGTGGGT
```

FIG. 3C-1

```
HTA7        GGGCGCGGCCACGATGTGCTCTGCGCTCTACGTGGGT
HTA18       GGGCGCGGCCACGATGTGCTCTGCGCTCTACGTGGGT
HTA20       GGGCGCGGCCACGATGTGCTCCGCGCTCTACGTGGGT
HTA22       GGGCGCGGCCACGATGTGCTCTGCGCTCTACGTGGGT
HTA26       GGGCGCGGCCACGATGTGCTCTGCGCTCTATGTGGGT
HTA35       GGGCGCGGCCACGATGTGCTCTGCGCTCTACGTGGGT
NZL-1       AGGCGCGGCCACGATGTGCTCTGCGCTCTACGTGGGT
HCV-1       CGGGAGCGCCACCCTCTGTTCGGCCCTCTACGTGGGG
HCV-J       TGGGGCGGCTGCTCTCTGTTCCGCTATGTACGTTGGG
HC-J6       GATGTCCGCCACGCTCTGCTCCGCTCTTTACGTGGGG
HC-J8       AATGGCAGCTACGGCCTGCTCGGCCTTGTATGTGGGA
                ** *  ** ** **  * ** ** **
```

FIG. 3C-2

```
HC-J6       -GCAGAAAGCGTCTAGCCATGGCGTTAGTATGAGTGTCGTACAGCCTCCAGGCCCCCCCC
HC-J8       -GCAGAAAGCGTCTAGCCATGGCGTTAGTATGAGTGTCGTACAGCCTCCAGGCCCCCCCC
HC-J4       -GCAGAAAGCGTCTAGCCATGGCGTTAGTATGAGTGTCGTGCAGCCTCCAGGACCCCCCC
HCV-1       -GCAGAAAGCGTCTAGCCATGGCGTTAGTATGAGTGTCGTGCAGCCTCCAGGACCCCCCC
NZL1        -GCGGAAAGCGCCTAGCCATGGCGTTAGTACGAGTGTCGTGCAGCCTCCAGGACCCCCCC
S83         -GCAGAAAGCGTCTAGCCATGGCGTTAGTATGAGTGTCGTACAGCCTCCAGGCCCCCCCC
HTA23       -GCAGAAAGCGTCTAGCCATGGCGTTAGTATGAGTGTCGTACAGCCTCCAGGCCCCCCCC
HTA30       -GCAGAAAGCGTCTAGCCATGGCGTTAGTATGAGTGTCGTACAGCCTCCAGGCCCCCCCC
HTA33       GCAGAAAGCGTCTAGCCATGGCGTTAGTATGAGTGTCGTACAGCCTCCAGGTCCCCCCC
             ** ******* ****************** ********* ********** *******

HC-J6       TCCCGGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTGCCGGGAAGAC
HC-J8       TCCCGGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTACCGGAAAGAC
HC-J4       TCCCGGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTGCCAGGACGAC
HCV-1       TCCCGGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTGCCAGGACGAC
NZL1        TCCCGGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATCGCTGGGGTGAC
S83         TCCCGGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTGCCGGGAAGAC
HTA23       TCCCGGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTGCCGGGAAGAC
HTA30       TCCCGGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTGCCAGGAAGAC
HTA33       TCCCGGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTGCCGGGAAGAC
            ************************************************  *  *  ***

HC-J6       TGGGTCCTTTCTTGGATAAACCCACTCTATGCCCGGTCATTTGGGCGTGCCCCCGCAAGA
HC-J8       TGGGTCCTTTCTTGGATAAACCCACTCTATGTCCGGTCATTTGGGCACGCCCCCGCAAGA
HC-J4       CGGGTCCTTTCTTGGATCAACCCGCTCAATGCCTGGAGATTTGGGCGTGCCCCCGCGAGA
HCV-1       CGGGTCCTTTCTTGGATCAACCCGCTCAATGCCTGGAGATTTGGGCGTGCCCCCGCAAGA
NZL1        CGGGTCCTTTCTTGGAGCAACCCGCTCAATACCCAGAAATTTGGGCGTGCCCCCGCGAGA
S83         TGGGTCCTTTCTTGGATAAACCCACTCTATGCCCGGCCATTTGGGCGTGCCCCCGCAAGA
HTA23       TGGGTCCTTTCTTGGATAAACCCACTCTATGCCCGGCCATTTGGGCGTGCCCCCGCAAGA
HTA30       TGGGTCCTTTCTTGGATAAACCCACTCTATGCCTGGCCATTTGGGCGTGCCCCCGCAAGA
HTA33       TGGGTCCTTTCTTGGATAAACCCACTCTATGCCCGGCCATTTGGGCGTGCCCCCGCAAGA
             ***************  ***** *** **  *  *  ********  ******** ***

HC-J6       CTGCTAGCCGAGTAG
HC-J8       CTGCTAGCCGAGTAG
HC-J4       CTGCTAGCCGAGTAG
HCV-1       CTGCTAGCCGAGTAG
NZL1        TCACTAGCCGAGTAG
S83         CTGCTAGCCGAGTAG
HTA23       CTGCTAGCCGAGTAG
HTA30       CTGCTAGCCGAGTAG
HTA33       CTGCTAGCCGAGTAG
            ************
```

FIG. 3D

LISTA DE SECUENCIAS

(1) INFORMACIÓN GENERAL

(i) SOLICITANTE: CHIRON CORPORATION

(ii) TÍTULO DE LA INVENCIÓN: ENSAYO DE RASTREO DE HETERODÚPLEX (HTA) PARA GENOTIPADO DEL HCV

(iii) NÚMERO DE SECUENCIAS: 52

(iv) DIRECCIÓN PARA LA CORRESPONDENCIA:

(A) DESTINATARIO: Chiron Corporation

(B) CALLE: 4560 Horton Street, R440

(C) CIUDAD: Emeryville

(D) ESTADO: California

(E) PAÍS: EE.UU.

(F) CÓDIGO POSTAL (ZIP): 94608-2916

(v) FORMA DE LECTURA POR ORDENADOR:

(A) TIPO DE MEDIO: disquete

(B) ORDENADOR: Compatible con PC de IBM

(C) SISTEMA OPERATIVO: PC-DOS/MS-DOS

(D) SOFTWARE: PatentIn Release nº 1.0, versión nº 1.30

(vi) DATOS DE LA SOLICITUD ACTUAL:

(A) NÚMERO DE SOLICITUD: no asignado

(B) FECHA DE PRESENTACIÓN: misma fecha que la presente

(C) CLASIFICACIÓN:

(viii) INFORMACIÓN SOBRE EL REPRESENTANTE/AGENTE:

(A) NOMBRE: Harbin, Alisa A.

(B) NÚMERO DE REGISTRO: 33.895

(C) NÚMERO DE REFERENCIA/EXPEDIENTE: 1226.100

(ix) INFORMACIÓN SOBRE TELECOMUNICACIONES:

(A) TELÉFONO: (510) 923-3274

(B) TELEFAX: (510) 655-3542

(C) TÉLEX: N/A

(2) INFORMACIÓN DE LA SEQ ID NO: 1:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 22 pares de bases

(B) TIPO: ácido nucleico

(C) CATENARIEDAD: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 1:

CCTGGTTGCT CTTTCTCTAT CT 22

(2) INFORMACIÓN DE LA SEQ ID NO: 2:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 21 pares de bases
(B) TIPO: ácido nucleico
(C) CATENARIEDAD: sencilla
(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 2:

GATGGCTTGT GGGATCCGGA G 21

(2) INFORMACIÓN DE LA SEQ ID NO: 3:

(i) CARACTERÍSTICAS DE LA SECUENCIA:
(A) LONGITUD: 21 pares de bases
(B) TIPO: ácido nucleico
(C) CATENARIEDAD: sencilla
(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 3:

GATGACCTCG GGGACGCGCA T 21

(2) INFORMACIÓN DE LA SEQ ID NO: 4:

(i) CARACTERÍSTICAS DE LA SECUENCIA:
(A) LONGITUD: 21 pares de bases
(B) TIPO: ácido nucleico
(C) CATENARIEDAD: sencilla
(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 4:

GACCAGTTCT GGAACACGAG C 21

(2) INFORMACIÓN DE LA SEQ ID NO: 5:

(i) CARACTERÍSTICAS DE LA SECUENCIA:
(A) LONGITUD: 21 pares de bases
(B) TIPO: ácido nucleico
(C) CATENARIEDAD: sencilla
(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 5:

CAAGGTCTGG GGTAAACGCA G 21

(2) INFORMACIÓN DE LA SEQ ID NO: 6:

(i) CARACTERÍSTICAS DE LA SECUENCIA:
(A) LONGITUD: 21 pares de bases

(B) TIPO: ácido nucleico
(C) CATENARIEDAD: sencilla
(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 6:

CCAGTTCATC ATCATATCCC A 21

(2) INFORMACIÓN DE LA SEQ ID NO: 7:

(i) CARACTERÍSTICAS DE LA SECUENCIA:
(A) LONGITUD: 22 pares de bases
(B) TIPO: ácido nucleico
(C) CATENARIEDAD: sencilla
(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 7:

TGGCCCTGCT CTCTTGCTTG AC 22

(2) INFORMACIÓN DE LA SEQ ID NO: 8:

(i) CARACTERÍSTICAS DE LA SECUENCIA:
(A) LONGITUD: 22 pares de bases
(B) TIPO: ácido nucleico
(C) CATENARIEDAD: sencilla
(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 8:

TTGCTCTTCT GTCGTGCGTC AC 22

(2) INFORMACIÓN DE LA SEQ ID NO: 9:

(i) CARACTERÍSTICAS DE LA SECUENCIA:
(A) LONGITUD: 22 pares de bases
(B) TIPO: ácido nucleico
(C) CATENARIEDAD: sencilla
(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 9:

TTGCTCTGTT CTCTTGCTTA AT 22

(2) INFORMACIÓN DE LA SEQ ID NO: 10:

(i) CARACTERÍSTICAS DE LA SECUENCIA:
(A) LONGITUD: 306 pares de bases
(B) TIPO: ácido nucleico

(C) CATENARIEDAD: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 10:

```
AACTCAAGCA TTGTGTATGA AGCGGCGGAC ATGATCATGC ACACCCCCGG GTGCGTGCCA      60
TGCGTCCGGG AGGGCAATCT CTCCCGCTGC TGGGTAGCGC TCACTCCCAC GCTCGCGGCC     120
AGAAACAGCA GCGTTCCTAC TACGACAATA CGACGCCATG TCGACTTGCT AGTAGGAGCG     180
GCTGCTTTTT GCTCCGCCAT GTACGTGGGG GACCTCTGCG GATCTATTTT CCTCGTCTCC     240
CAACTGTTCA CCTTCTCGCC CCGCCGGCAT CATACAGTAC AGGACTGCAA TTGCTCGATC     300
TATCCC                                                               306
```

(2) INFORMACIÓN DE LA SEQ ID NO: 11:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 306 pares de bases

(B) TIPO: ácido nucleico

(C) CATENARIEDAD: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 11:

```
AACTCAAGCA TCGTGTATGA GGCAGCGGAA GTGATCATGC ACATTCCCGG GTGCGTGCCC      60
TGCGTTCGGG AGAGCAATCT CTCCCGCTGC TGGGTAGCGC TCACCCCCAC ACTCGCGGCC     120
AGGAACAGCA GCGTCCCCAC CACGACAATA CGACGCCACG TCGACTTGCT CGTTGGGGCG     180
GCTGCCTTCT GCTCCGCTAT GTATGTGGGG GATCTCTGCG GATCTGTTTT CCTTGTCTCC     240
CAACTGTTCA CCTTTTCGCC TCGCCGGCAT GAGACAGTAC AGGACTGCAA TTGTTCAATC     300
TATCCC                                                               306
```

(2) INFORMACIÓN DE LA SEQ ID NO: 12:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 306 pares de bases

(B) TIPO: ácido nucleico

(C) CATENARIEDAD: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 12:

```
AACTCAAGCA TAGTATATGA GGCAGCGGAC ATAATCATGC ATACCCCCGG GTGCGTGCCC      60
TGTGTTCGGG AGGTCAACTC CTCCCGCTGC TGGGCAGCGC TCACCCCTAC GCTCGCGGCC     120
AGGAACTCCA GCGTGCCCAC TACGACAATA CGACGCCACG TCGACTTGCT CGTTGGGGCG     180
```

GCTGCTTTCT GCTCCGCTAT GTACGTGGGG GATCTATGCG GATCTGTTCT ACTTGTCTCT 240

CAGCTGTTCA CCTTCTCACC TCGCCGGCAC GAGACAGTGC AGGACTGCAA TTGTTCAATC 300

TATCCC 306

(2) INFORMACIÓN DE LA SEQ ID NO: 13:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 306 pares de bases

(B) TIPO: ácido nucleico

(C) CATENARIEDAD: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 13:

AACACGAGCA TTGTGTATGA GGCAGCGGAC TTGATCATGC ACGTCCCCGG GTGCGTGCCC 60

TGCGTTCGGG AGGGCAACTC CTCCCGATGC TGGGTAGCGC TCACTCCCAC GATCGCGGCC 120

AGGAACAGCA GTGTCCCCGT TACGACCATA CGACGCCACG TCGATTTGCT CGTTGGGGCG 180

GCTGCTCTTT GCTCCGCCAT GTACGTGGGG GATCTCTGCG GATCTGTCTT CCTCGCTTCC 240

CAGTTGTTCA CTTTCTCGCC TCGCCAGCAT CAGACGGTAC AGGACTGCAA CTGCTCAATC 300

TATCCC 306

(2) INFORMACIÓN DE LA SEQ ID NO: 14:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 306 pares de bases

(B) TIPO: ácido nucleico

(C) CATENARIEDAD: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 14:

AACTCAAGCA TCGTGTATGA GGCGGCGGAA GTGATCATGC ACATTCCTGG GTGCGTGCCC 60

TGCGTTCGGG AGGGCGACTT CTCCCGCTGC TGGGTAGCGC TCACCCCCAC ACTCGCGGCC 120

AGGAATAACA GCGTCCCCAC TACGACAATA CGACGCCACG TCGACTTGCT CGTTGGGGCG 180

GCTGCCTTCT GCTCCGCTAT GTACGTGGGG GATCTCTGCG GATCTGTTTT CCTTGTCTCC 240

CAACTGTTCA CCTTTTCGCC TCGCCGGCAT GCGACAGTAC AGGACTGCAA TTGTTCAATC 300

TATCCC 306

(2) INFORMACIÓN DE LA SEQ ID NO: 15:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 306 pares de bases

(B) TIPO: ácido nucleico

(C) CATENARIEDAD: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 15:

```
AACTCGAGTA TTGTGTACGA GGCGGCCGAT GCCATCCTGC ACACTCCGGG GTGCGTCCCT      60
TGCGTTCGTG AGGGCAACGC CTCGAGGTGT TGGGTGGCGA TGACCCCTAC GGTGGCCACC     120
AGGGATGGCA AACTCCCCGC GACGCAGCTT CGACGTCACA TCGATCTGCT TGTCGGGAGC     180
GCCACCCTCT GTTCGGCCCT CTACGTGGGG GACCTATGCG GGTCTGTCTT TCTTGTCGGC     240
CAACTGTTCA CCTTCTCTCC CAGGCGCCAC TGGACGACGC AAGGTTGCAA TTGCTCTATC     300
TATCCC                                                                306
```

(2) INFORMACIÓN DE LA SEQ ID NO: 16:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 306 pares de bases

(B) TIPO: ácido nucleico

(C) CATENARIEDAD: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 16:

```
AACTCAAGTA TTGTGTATGA GGCAGCGGAC ATGATCATGC ACACCCCCGG GTGCGTGCCC      60
TGCGTCCGGG AGAGTAATTT CTCCCGTTGC TGGGTAGCGC TCACTCCCAC GCTCGCGGCC     120
AGGAACAGCA GCATCCCCAC CACGACAATA CGACGCCACG TCGATTTGCT CGTTGGGGCG     180
GCTGCTCTCT GTTCCGCTAT GTACGTTGGG GATCTCTGCG GATCCGTTTT TCTCGTCTCC     240
CAGCTGTTCA CCTTCTCACC TCGCCGGTAT GAGACGGTAC AAGATTGCAA TTGCTCAATC     300
TATCCC                                                                306
```

(2) INFORMACIÓN DE LA SEQ ID NO: 17:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 306 pares de bases

(B) TIPO: ácido nucleico

(C) CATENARIEDAD: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 17:

```
AATGATAGCA TTACCTGGCA ACTCCAGGCT GCTGTCCTCC ACGTCCCCGG GTGCGTCCCG      60
```

TGCGAGAAAG TGGGGAATAC ATCTCGGTGC TGGATACCGG TCTCACCGAA TGTGGCCGTG 120

CAGCAGCCCG GCGCCCTCAC GCAGGGCTTA CGGACGCACA TTGACATGGT TGTGATGTCC 180

GCCACGCTCT GCTCCGCTCT TTACGTGGGG GACCTCTGCG GTGGGGTGAT GCTTGCAGCC 240

CAGATGTTCA TTGTCTCGCC ACAGCACCAC TGGTTTGTGC AAGACTGCAA TTGCTCCATC 300

TACCCT 306

(2) INFORMACIÓN DE LA SEQ ID NO: 18:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 306 pares de bases

(B) TIPO: ácido nucleico

(C) CATENARIEDAD: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 18:

AACAACAGCA TCACCTGGCA GCTCACTGAC GCAGTTCTCC ATCTTCCTGG ATGCGTCCCA 60

TGTGAGAATG ATAATGGCAC CTTGCATTGC TGGATACAAG TAACACCCAA CGTGGCTGTG 120

AAACACCGCG GTGCGCTCAC TCGTAGCCTG CGAACACACG TCGACATGAT CGTAATGGCA 180

GCTACGGCCT GCTCGGCCTT GTATGTGGGA GATGTGTGCG GGGCCGTGAT GATTCTATCG 240

CAGGCTTTCA TGGTATCACC ACAACGCCAC AACTTCACCC AAGAGTGCAA CTGTTCCATC 300

TACCAA 306

(2) INFORMACIÓN DE LA SEQ ID NO: 19:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 306 pares de bases

(B) TIPO: ácido nucleico

(C) CATENARIEDAD: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 19:

AATAGCAGTA TTGTGTATGA GGCCGATGAT GTCATTCTGC ACACACCCGG CTGTGTACCT 60

TGTGTCCAGG ACGGCAATAC ATCTACGTGC TGGACCCCAG TGACACCTAC AGTGGCAGTC 120

AGGTACGTCG GAGCAACTAC TGCTTCGATA CGCAGTCATG TGGACCTATT AGTAGGCGCG 180

GCCACGATGT GCTCTGCGCT CTACGTGGGT GATATGTGTG GGGCTGTCTT TCTCGTGGGA 240

CAAGCCTTCA CGTTCAGACC TCGACGCCAT CAAACGGTCC AGACCTGTAA CTGCTCGCTG 300

TACCCA 306

(2) INFORMACIÓN DE LA SEQ ID NO: 20:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 333 pares de bases

(B) TIPO: ácido nucleico

(C) CATENARIEDAD: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 20:

```
CGCAACTCCA CGGGGCTTTA CCACGTCACC AATGATTGCC CTAACTCGAG TATTGTGTAC      60
GAGACGGCCG ATGCCATCCT GCACACTCCG GGGTGCGTCC CTTGTGTTCG CGAGGGCAAC     120
GCCTCGAGGT GTTGGGTGGC GATGACCCCT ACGGTGGCCA CCAGGGATGG CAAACTCCCC     180
GCGACGCAGC TTCGACGTCA CATCGATCTG CTTGTCGGGA GCGCCACCCT CTGTTCGGCC     240
CTCTACGTGG GGGATCTGTG CGGGTCTGTC TTTCTTGTCG GCCAACTGTT TACCTTCTCT     300
CCCAGGCGCC ACTGGACGAC GCAAGGTTGC AAT                                  333
```

(2) INFORMACIÓN DE LA SEQ ID NO: 21:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 333 pares de bases

(B) TIPO: ácido nucleico

(C) CATENARIEDAD: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 21:

```
AAGAACACCA GCGACAGCTA CATGGTGACC AATGACTGCC AAAATGACAG CATCACCTGG      60
CAGCTTGAGG CTGCGGTCCT CCACGTCCCC GGGTGCGTCC CGTGCGAGAG AGTGGGAAAT     120
ACATCTCGGT GCTGGATACC GGTCTCACCA AACGTGGCTG TGCGGCAGCC CGGCGCCCTC     180
ACGCAGGGCT TGCGGACGCA CATCGACATG ATTGTGATGT CCGCCACGCT CTGCTCCGCT     240
CTCTACGTGG GGGACCTCTG TGGCGGGATG ATGCTCGCAG CCCAGATGTT CATCGTTTCG     300
CCGCAGAACC ACTGGTTCGT GCAGGAATGC AAT                                  333
```

(2) INFORMACIÓN DE LA SEQ ID NO: 22:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 333 pares de bases

(B) TIPO: ácido nucleico

(C) CATENARIEDAD: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 22:

```
AGGAACATCA GTTCTAGCTA CTACGCCACT AATGACTGCT CGAACAACAG CATCACCTGG     60
CAGCTCACCA ACGCAGTTCT CCACCTTCCC GGATGCGTCC CATGTGAGAA TAATAATGGC    120
ACCTTGCATT GCTGGATACA AGTAACACCT AATGTGGCCG TAAAACATCG CGGCGCACTC    180
ACTCACAACC TGCGGACACA TGTCGACATG ATCGTAATGG CAGCTACGGT CTGTTCGGCC    240
TTGTACGTAG GAGACGTGTG TGGGGCTGTG ATGATTGTGT CTCAGGCCCT TATAATATCA    300
CCAGAACACC ATAACTTCAC CCAAGAGTGC AAC                                 333
```

(2) INFORMACIÓN DE LA SEQ ID NO: 23:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 333 pares de bases

(B) TIPO: ácido nucleico

(C) CATENARIEDAD: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 23:

```
AAGGACACCG GCGACTCCTA CATGCCGACC AACGATTGCT CCAACTCTAG TATCGTTTGG     60
CAGCTTGAAG GAGCAGTGCT TCATACTCCT GGATGCGTCC CTTGTGAGCG TACCGCCAAC    120
GTCTCTCGAT GTTGGGTGCC GGTTGCCCCC AATCTCGCCA TAAGTCAACC TGGCGCTCTC    180
ACTAAGGGCC TGCGAGCACA CATCGATATC ATCGTGATGT CTGCTACGGT CTGTTCTGCC    240
CTTTATGTGG GGGACGTGTG TGGCGCGCTG ATGCTGGCCG CTCAGGTCGT CGTCGTGTCG    300
CCACAACACC ATACGTTTGT CCAGGAATGC AAC                                 333
```

(2) INFORMACIÓN DE LA SEQ ID NO: 24:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 333 pares de bases

(B) TIPO: ácido nucleico

(C) CATENARIEDAD: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 24:

```
AAAAACACCA GCATCTCCTA TATGGCGACC AACGACTGCT CCAATTCCAG CATCGCTTGG     60
CAGTTTGACG GCGCAGTGCT CCATACTCCT GGATGTGTCC CTTGCGAACG GACCGGCAAC    120
GCGTCCCGGT GTTGGGTGCC GGTTGCCCCC AATGTGGCTA TAAGACAACC CGGCGCCCTC    180
ACTAAGGGCA TACGAACGCA CATTGATGTC ATCGTAATGT CTGCTACGCT CTGTTCTGCC    240
CTTTACGTGG GGGACGTGTG TGGTGCGCTG ATGATTGCCG CTCAGGTCGT CATTGTGTCT    300
CCGCAGCATC ACCACTTTGT CCAGGACTGC AAT                                 333
```

(2) INFORMACIÓN DE LA SEQ ID NO: 25:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 333 pares de bases

(B) TIPO: ácido nucleico

(C) CATENARIEDAD: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 25:

```
AAGAACACCA GCGACTCCTA CATGGCGACT AACGACTGCT CTAACTCCAG CATCGTTTGG      60
CAGCTTGAGG ACGCAGTGCT CCATGTCCCT GGATGTGTCC CTTGTGAGAA GACTGGCAAT     120
ACGTCTCGGT GCTGGGTGCC GGTTACCCCC AATGTGGCTA CAAGTCAACC CGGCGCTCTC     180
ACCAGGGGCT TGCGGACGCA CATCGATGTC ATCGTGATGT CAGCCACGCT CTGCTCCGCT     240
CTCTATGTGG GGGACGTGTG TGGCGCGTTG ACGATAGCCG CTCAGGTTGT CATCGTATCG     300
CCACGGCACC ACCACTTTGT CCAGGACTGC AAT                                  333
```

(2) INFORMACIÓN DE LA SEQ ID NO: 26:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 333 pares de bases

(B) TIPO: ácido nucleico

(C) CATENARIEDAD: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 26:

```
AAGAACACCA GCACCTCCTA CATGGTGACT AACGATTGCT CCAACTCCAG CATCGTTTGG      60
CAACTTGAAG GCGCAGTGCT CCATGTTCCT GGATGTGTCC CTTGTGAGCA GATCGGCAAC     120
GTGTCTCAGT GTTGGGTGCC GGTTACCCCC AATATGGCCA TAAGTACACC CGGCGCTCTC     180
ACTAAGGGCT TGCGAACGCA CATCGACGGC ATCGTGATGT CCGCTACGCT CTGTTCTGCC     240
CTTTATGTGG GGGACGTGTG TGGCGCGTTG ATGATAGCCG CCCAGGTCGT CATCGTATCG     300
CCACAGCACC ACCACTTTGT CCACGACTGC AAC                                  333
```

(2) INFORMACIÓN DE LA SEQ ID NO: 27:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 333 pares de bases

(B) TIPO: ácido nucleico

(C) CATENARIEDAD: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 27:

```
CGCAACTCCA CGGGGCTTTA CCACGTCACC AATGATTGCC CTAACTCGAG TATTGTGTAC    60

GAGGCGGCCG ATGCCATCCT GCACACTCCG GGGTGCGTCC CTTGCGTTCG TGAGGGCAAC   120

GCCTCGAGGT GTTGGGTGGC GATGACCCCT ACGGTGGCCA CCAGGGATGG CAAACTCCCC   180

GCGACGCAGC TTCGACGTCA CATCGATCTG CTTGTCGGGA GCGCCACCCT CTGTTCGGCC   240

CTCTACGTGG GGGACCTATG CGGGTCTGTC TTTCTTGTCG GCCAACTGTT CACCTTCTCT   300

CCCAGGCGCC ACTGGACGAC GCAAGGTTGC AAT                                 333
```

(2) INFORMACIÓN DE LA SEQ ID NO: 28:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 333 pares de bases

(B) TIPO: ácido nucleico

(C) CATENARIEDAD: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 28:

```
CGCAACGTGT CCGGGATATA CCATGTCACG AACGACTGCT CCAACTCAAG TATTGTGTAT    60

GAGGCAGCGG ACATGATCAT GCACACCCCC GGGTGCGTGC CCTGCGTCCG GGAGAGTAAT   120

TTCTCCCGTT GCTGGGTAGC GCTCACTCCC ACGCTCGCGG CCAGGAACAG CAGCATCCCC   180

ACCACGACAA TACGACGCCA CGTCGATTTG CTCGTTGGGG CGGCTGCTCT CTGTTCCGCT   240

ATGTACGTTG GGGATCTCTG CGGATCCGTT TTTCTCGTCT CCCAGCTGTT CACCTTCTCA   300

CCTCGCCGGT ATGAGACGGT ACAAGATTGC AAT                                 333
```

(2) INFORMACIÓN DE LA SEQ ID NO: 29:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 333 pares de bases

(B) TIPO: ácido nucleico

(C) CATENARIEDAD: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 29:

```
AGGAACATTA GTTCTAGCTA CTACGCCACT AATGATTGCT CAAACAACAG CATCACCTGG    60

CAGCTCACTG ACGCAGTTCT CCATCTTCCT GGATGCGTCC CATGTGAGAA TGATAATGGC   120

ACCTTGCATT GCTGGATACA AGTAACACCC AACGTGGCTG TGAAACACCG CGGTGCGCTC   180

ACTCGTAGCC TGCGAACACA CGTCGACATG ATCGTAATGG CAGCTACGGC CTGCTCGGCC   240

TTGTATGTGG GAGATGTGTG CGGGGCCGTG ATGATTCTAT CGCAGGCTTT CATGGTATCA   300

CCACAACGCC ACAACTTCAC CCAAGAGTGC AAC                                 333
```

(2) INFORMACIÓN DE LA SEQ ID NO: 30:

(i) CARACTERÍSTICAS DE LA SECUENCIA:
- (A) LONGITUD: 333 pares de bases
- (B) TIPO: ácido nucleico
- (C) CATENARIEDAD: sencilla
- (D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 30:

```
CGGAATACGT CTGGCCTCTA CGTCCTTACC AACGACTGTT CCAATAGCAG TATTGTGTAT      60
GAGGCCGATG ATGTCATTCT GCACACACCC GGCTGTGTAC CTTGTGTCCA GGACGGCAAT     120
ACATCTACGT GCTGGACCCC AGTGACACCT ACAGTGGCAG TCAGGTACGT CGGAGCAACT     180
ACTGCTTCGA TACGCAGTCA TGTGGACCTA TTAGTAGGCG CGGCCACGAT GTGCTCTGCG     240
CTCTACGTGG GTGATATGTG TGGGGCTGTC TTTCTCGTGG GACAAGCCTT CACGTTCAGA     300
CCTCGACGCC ATCAAACGGT CCAGACCTGT AAC                                  333
```

(2) INFORMACIÓN DE LA SEQ ID NO: 31:

(i) CARACTERÍSTICAS DE LA SECUENCIA:
- (A) LONGITUD: 333 pares de bases
- (B) TIPO: ácido nucleico
- (C) CATENARIEDAD: sencilla
- (D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 31:

```
AAGAACATCA GTACCGGCTA CATGGTGACC AACGACTGCA CCAATGATAG CATTACCTGG      60
CAACTCCAGG CTGCTGTCCT CCACGTCCCC GGGTGCGTCC CGTGCGAGAA AGTGGGGAAT     120
ACATCTCGGT GCTGGATACC GGTCTCACCG AATGTGGCCG TGCAGCAGCC CGGCGCCCTC     180
ACGCAGGGCT TACGGACGCA CATTGACATG GTTGTGATGT CCGCCACGCT CTGCTCCGCT     240
CTTTACGTGG GGGACCTCTG CGGTGGGGTG ATGCTTGCAG CCCAGATGTT CATTGTCTCG     300
CCACAGCACC ACTGGTTTGT GCAAGACTGC AAT                                  333
```

(2) INFORMACIÓN DE LA SEQ ID NO: 32:

(i) CARACTERÍSTICAS DE LA SECUENCIA:
- (A) LONGITUD: 277 pares de bases
- (B) TIPO: ácido nucleico
- (C) CATENARIEDAD: sencilla
- (D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 32:

TCATCCAACA TCTAGTCTAG AGTGGCGGAA TACGTCTGGC CTCTATGTCC TTACCAACGA 60

CTGTTCCAAT AACATTATTG TGTATGAGGC CGATGACGTC ATCCTGCACA CGCCCGGCTG 120

TGTACCTTGT GTTCAGGACG GTAATACATC CAAGTGCTGG ACCCCAGTGA CACCTACAGT 180

GGCAGTCAGG TACGTCGGAG CAACCACCGC TTCAATACGC AGCCACGTGG ACCTATTATT 240

GGGCGCGGCC ACGATGTGCT CTGCGCTCTA CGTGGGT 277

(2) INFORMACIÓN DE LA SEQ ID NO: 33:

(i) CARACTERÍSTICAS DE LA SECUENCIA:
- (A) LONGITUD: 277 pares de bases
- (B) TIPO: ácido nucleico
- (C) CATENARIEDAD: sencilla
- (D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 33:

TCATCCAACA TCTAGTCTAG AGTGGCGGAA TACGTCTGGC CTCTATGTCC TTACCAACGA 60

CTGTTCCAAT AACATCATTG TGTATGAGGC CGATGACGTC ATCCTGCACG CACCCGGCTG 120

TGTACCTTGT GTTCAGGACG GCAATACATC CACGTGCTGG ACCCCAGTGA CACCTACAGT 180

GGCAGTCAGG TACGTCGGAG CAACCACCGC TTCAATACGC AGCCATGTGG ACCTATTAGT 240

GGGCGCGGCC ACGATGTGCT CTGCGCTCTA CGTGGGT 277

(2) INFORMACIÓN DE LA SEQ ID NO: 34:

(i) CARACTERÍSTICAS DE LA SECUENCIA:
- (A) LONGITUD: 277 pares de bases
- (B) TIPO: ácido nucleico
- (C) CATENARIEDAD: sencilla
- (D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 34:

TCATCCAACA TCTAGTCTAG AGTGGCGGAA TACGTCTGGC CTCTATGTCC TTACCAACGA 60

CTGTTCCAAT AATATTATTG TGTATGAGGC CGACGACGTC ATCCTGCACG CCCCCGGCTG 120

TGTACCTTGT GTTCAGGACG GCAATACATC CACGTGCTGG ATCCCAGTGA CACCTACAGT 180

GGCAGTCAGG TACGCCGGAG CAACCACCGC TTCAATACGC AGCCATGTGG ACCTGTTAGT 240

GGGCGCGGCC ACGATGTGCT CTGCGCTCTA CGTGGGT 277

(2) INFORMACIÓN DE LA SEQ ID NO: 35:

(i) CARACTERÍSTICAS DE LA SECUENCIA:
- (A) LONGITUD: 277 pares de bases
- (B) TIPO: ácido nucleico

(C) CATENARIEDAD: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 35:

```
TCATCCAACA TCTAGTCTAG AGTGGCGGAA TACGTCTGGC CTCTATGTCC TTACCAACGA      60
CTGTTCCAAT AACATTATTG TGTATGAGGC CGATGACGTC ATCCTGCACA CACCCGGCTG     120
TGTACCTTGT GTTCAGGACG GCAATACATC CACGTGCTGG ACCCCAGTGA CACCTACAGT     180
ATCAGTCAGG TACGTCGGAG CAACCACCGC TTCAATACGC AGCCATGTGG ACCTACTATT     240
GGGCGCGGCC ACGATGTGCT CCGCGCTCTA CGTGGGT                              277
```

(2) INFORMACIÓN DE LA SEQ ID NO: 36:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 277 pares de bases

(B) TIPO: ácido nucleico

(C) CATENARIEDAD: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 36:

```
TCATCCAACA TCTAGTCTAG AGTGGCGGAA TACGTCTGGC CTCTATGTCC TTACCAACGA      60
CTGTTCCAAT AACAGTATTG TGTATGAGGC CGATCACGTC ATCCTGCACA CACCCGGCTG     120
TGTACCTTGT GTTCAAGCCA ACAATAAATC CAAATGCTGG ACCCCAGTGA CACCTACAGT     180
ATCAGTCGAG TACGTCGGAG CAACCACCGC TTCAATACGC AGCCATGTGG ACCTACTATT     240
GGGCGCGGCC ACGATGTGCT CTGCGCTCTA CGTGGGT                              277
```

(2) INFORMACIÓN DE LA SEQ ID NO: 37:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 277 pares de bases

(B) TIPO: ácido nucleico

(C) CATENARIEDAD: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 37:

```
TCATCCAACA TCTAGTCTAG AGTGGCGGAA TACGTCTGGC CTCTATGTCC TTACCAACGA      60
CTGTTCTAAT AACATTATTG TGTATGAGGC CGATGACGTC ATCCTGCACA CACCCGGCTG     120
TGTACCTTGT GTTCAGGACG GCAATGCATC CACGTGCTGG ACCCCAGTAA CACCTACAGT     180
ATCAGTCAGG TACGTCGGAG CAACCACCGC TTCAGTACGC AGCCATGTGG ACCTACTATT     240
GGGCGCGGCC ACGATGTGCT CTGCGCTCTA TGTGGGT                              277
```

(2) INFORMACIÓN DE LA SEQ ID NO: 38:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 277 pares de bases

(B) TIPO: ácido nucleico

(C) CATENARIEDAD: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 38:

```
TCATCCAACA TCTAGTCTAG AGTGGCGGAA TACGTCTGGC CTCTATGTCC TCACCAACGA      60

CTGTTCCAAC AACATTATTG TGTATGAGGC CGATGACGTC ATTCTGCACA CGCCCGGCTG     120

CGTACCTTGT GTACAGGACG GCAATACATC CACGTGCTGG ACCCCAGTGA CACCTACAGT     180

GGCAGTCAGG TACGTCGGAG CAACTACCGC TTCAATACGC AGCCATGTGG ACCTATTATT     240

GGGCGCGGCC ACGATGTGCT CTGCGCTCTA CGTGGGT                              277
```

(2) INFORMACIÓN DE LA SEQ ID NO: 39:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 277 pares de bases

(B) TIPO: ácido nucleico

(C) CATENARIEDAD: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 39:

```
TCATCCAGCA GCCAGTCTAG AGTGGCGGAA TACGTCTGGC CTCTACGTCC TTACCAACGA      60

CTGTTCCAAT AGCAGTATTG TGTATGAGGC CGATGATGTC ATTCTGCACA CACCCGGCTG     120

TGTACCTTGT GTCCAGGACG GCAATACATC TACGTGCTGG ACCCCAGTGA CACCTACAGT     180

GGCAGTCAGG TACGTCGGAG CAACTACTGC TTCGATACGC AGTCATGTGG ACCTATTAGT     240

AGGCGCGGCC ACGATGTGCT CTGCGCTCTA CGTGGGT                              277
```

(2) INFORMACIÓN DE LA SEQ ID NO: 40:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 277 pares de bases

(B) TIPO: ácido nucleico

(C) CATENARIEDAD: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 40:

TGTGCCCGCT TCGGCCTACC AAGTGCGCAA CTCCACGGGG CTTTACCACG TCACCAATGA 60

TTGCCCTAAC TCGAGTATTG TGTACGAGGC GGCCGATGCC ATCCTGCACA CTCCGGGGTG 120

CGTCCCTTGC GTTCGTGAGG GCAACGCCTC GAGGTGTTGG GTGGCGATGA CCCCTACGGT 180

GGCCACCAGG GATGGCAAAC TCCCCGCGAC GCAGCTTCGA CGTCACATCG ATCTGCTTGT 240

CGGGAGCGCC ACCCTCTGTT CGGCCCTCTA CGTGGGG 277

(2) INFORMACIÓN DE LA SEQ ID NO: 41:

(i) CARACTERÍSTICAS DE LA SECUENCIA:
- (A) LONGITUD: 277 pares de bases
- (B) TIPO: ácido nucleico
- (C) CATENARIEDAD: sencilla
- (D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 41:

CATCCCAGCT TCCGCTTACG AGGTGCGCAA CGTGTCCGGG ATATACCATG TCACGAACGA 60

CTGCTCCAAC TCAAGTATTG TGTATGAGGC AGCGGACATG ATCATGCACA CCCCCGGGTG 120

CGTGCCCTGC GTCCGGGAGA GTAATTTCTC CCGTTGCTGG GTAGCGCTCA CTCCCACGCT 180

CGCGGCCAGG AACAGCAGCA TCCCCACCAC GACAATACGA CGCCACGTCG ATTTGCTCGT 240

TGGGGCGGCT GCTCTCTGTT CCGCTATGTA CGTTGGG 277

(2) INFORMACIÓN DE LA SEQ ID NO: 42:

(i) CARACTERÍSTICAS DE LA SECUENCIA:
- (A) LONGITUD: 277 pares de bases
- (B) TIPO: ácido nucleico
- (C) CATENARIEDAD: sencilla
- (D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 42:

CACCCCGGTC TCCGCTGCCG AAGTGAAGAA CATCAGTACC GGCTACATGG TGACCAACGA 60

CTGCACCAAT GATAGCATTA CCTGGCAACT CCAGGCTGCT GTCCTCCACG TCCCCGGGTG 120

CGTCCCGTGC GAGAAAGTGG GGAATACATC TCGGTGCTGG ATACCGGTCT CACCGAATGT 180

GGCCGTGCAG CAGCCCGGCG CCCTCACGCA GGGCTTACGG ACGCACATTG ACATGGTTGT 240

GATGTCCGCC ACGCTCTGCT CCGCTCTTTA CGTGGGG 277

(2) INFORMACIÓN DE LA SEQ ID NO: 43:

(i) CARACTERÍSTICAS DE LA SECUENCIA:
- (A) LONGITUD: 277 pares de bases
- (B) TIPO: ácido nucleico

(C) CATENARIEDAD: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 43:

```
AGTGCCAGTG TCTGCAGTGG AAGTCAGGAA CATTAGTTCT AGCTACTACG CCACTAATGA      60
TTGCTCAAAC AACAGCATCA CCTGGCAGCT CACTGACGCA GTTCTCCATC TTCCTGGATG     120
CGTCCCATGT GAGAATGATA ATGGCACCTT GCATTGCTGG ATACAAGTAA CACCCAACGT     180
GGCTGTGAAA CACCGCGGTG CGCTCACTCG TAGCCTGCGA ACACACGTCG ACATGATCGT     240
AATGGCAGCT ACGGCCTGCT CGGCCTTGTA TGTGGGA                               277
```

(2) INFORMACIÓN DE LA SEQ ID NO: 44:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 194 pares de bases

(B) TIPO: ácido nucleico

(C) CATENARIEDAD: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 44:

```
GCAGAAAGCG TCTAGCCATG GCGTTAGTAT GAGTGTCGTA CAGCCTCCAG GCCCCCCCCT      60
CCCGGGAGAG CCATAGTGGT CTGCGGAACC GGTGAGTACA CCGGAATTGC CGGGAAGACT     120
GGGTCCTTTC TTGGATAAAC CCACTCTATG CCCGGTCATT TGGGCGTGCC CCCGCAAGAC     180
TGCTAGCCGA GTAG                                                       194
```

(2) INFORMACIÓN DE LA SEQ ID NO: 45:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 194 pares de bases

(B) TIPO: ácido nucleico

(C) CATENARIEDAD: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 45:

```
GCAGAAAGCG TCTAGCCATG GCGTTAGTAT GAGTGTCGTA CAGCCTCCAG GCCCCCCCCT      60
CCCGGGAGAG CCATAGTGGT CTGCGGAACC GGTGAGTACA CCGGAATTAC CGGAAAGACT     120
GGGTCCTTTC TTGGATAAAC CCACTCTATG TCCGGTCATT TGGGCACGCC CCCGCAAGAC     180
TGCTAGCCGA GTAG                                                       194
```

(2) INFORMACIÓN DE LA SEQ ID NO: 46:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 194 pares de bases
(B) TIPO: ácido nucleico
(C) CATENARIEDAD: sencilla
(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 46:

```
GCAGAAAGCG TCTAGCCATG GCGTTAGTAT GAGTGTCGTG CAGCCTCCAG GACCCCCCCT   60
CCCGGGAGAG CCATAGTGGT CTGCGGAACC GGTGAGTACA CCGGAATTGC CAGGACGACC  120
GGGTCCTTTC TTGGATCAAC CCGCTCAATG CCTGGAGATT TGGGCGTGCC CCCGCGAGAC  180
TGCTAGCCGA GTAG                                                    194
```

(2) INFORMACIÓN DE LA SEQ ID NO: 47:

(i) CARACTERÍSTICAS DE LA SECUENCIA:
(A) LONGITUD: 194 pares de bases
(B) TIPO: ácido nucleico
(C) CATENARIEDAD: sencilla
(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 47:

```
GCAGAAAGCG TCTAGCCATG GCGTTAGTAT GAGTGTCGTG CAGCCTCCAG GACCCCCCCT   60
CCCGGGAGAG CCATAGTGGT CTGCGGAACC GGTGAGTACA CCGGAATTGC CAGGACGACC  120
GGGTCCTTTC TTGGATCAAC CCGCTCAATG CCTGGAGATT TGGGCGTGCC CCCGCAAGAC  180
TGCTAGCCGA GTAG                                                    194
```

(2) INFORMACIÓN DE LA SEQ ID NO: 48:

(i) CARACTERÍSTICAS DE LA SECUENCIA:
(A) LONGITUD: 194 pares de bases
(B) TIPO: ácido nucleico
(C) CATENARIEDAD: sencilla
(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 48:

```
GCGGAAAGCG CCTAGCCATG GCGTTAGTAC GAGTGTCGTG CAGCCTCCAG GACCCCCCCT   60
CCCGGGAGAG CCATAGTGGT CTGCGGAACC GGTGAGTACA CCGGAATCGC TGGGGTGACC  120
GGGTCCTTTC TTGGAGCAAC CCGCTCAATA CCCAGAAATT TGGGCGTGCC CCCGCGAGAT  180
CACTAGCCGA GTAG                                                    194
```

(2) INFORMACIÓN DE LA SEQ ID NO: 49:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 194 pares de bases

(B) TIPO: ácido nucleico

(C) CATENARIEDAD: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 49:

```
GCAGAAAGCG TCTAGCCATG GCGTTAGTAT GAGTGTCGTA CAGCCTCCAG GCCCCCCCCT      60
CCCGGGAGAG CCATAGTGGT CTGCGGAACC GGTGAGTACA CCGGAATTGC CGGGAAGACT     120
GGGTCCTTTC TTGGATAAAC CCACTCTATG CCCGGCCATT TGGGCGTGCC CCCGCAAGAC     180
TGCTAGCCGA GTAG                                                      194
```

(2) INFORMACIÓN DE LA SEQ ID NO: 50:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 194 pares de bases

(B) TIPO: ácido nucleico

(C) CATENARIEDAD: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 50:

```
GCAGAAAGCG TCTAGCCATG GCGTTAGTAT GAGTGTCGTA CAGCCTCCAG GCCCCCCCCT      60
CCCGGGAGAG CCATAGTGGT CTGCGGAACC GGTGAGTACA CCGGAATTGC CGGGAAGACT     120
GGGTCCTTTC TTGGATAAAC CCACTCTATG CCCGGCCATT TGGGCGTGCC CCCGCAAGAC     180
TGCTAGCCGA GTAG                                                      194
```

(2) INFORMACIÓN DE LA SEQ ID NO: 51:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 194 pares de bases

(B) TIPO: ácido nucleico

(C) CATENARIEDAD: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 51:

```
GCAGAAAGCG TCTAGCCATG GCGTTAGTAT GAGTGTCGTA CAGCCTCCAG GCCCCCCCCT      60
CCCGGGAGAG CCATAGTGGT CTGCGGAACC GGTGAGTACA CCGGAATTGC CAGGAAGACT     120
GGGTCCTTTC TTGGATAAAC CCACTCTATG CCTGGCCATT TGGGCGTGCC CCCGCAAGAC     180
TGCTAGCCGA GTAG                                                      194
```

(2) INFORMACIÓN DE LA SEQ ID NO: 52:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 194 pares de bases

(B) TIPO: ácido nucleico

(C) CATENARIEDAD: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 52:

```
GCAGAAAGCG TCTAGCCATG GCGTTAGTAT GAGTGTCGTA CAGCCTCCAG GTCCCCCCCT      60

CCCGGGAGAG CCATAGTGGT CTGCGGAACC GGTGAGTACA CCGGAATTGC CGGGAAGACT     120

GGGTCCTTTC TTGGATAAAC CCACTCTATG CCCGGCCATT TGGGCGTGCC CCCGCAAGAC     180

TGCTAGCCGA GTAG                                                       194
```