

(12) SOLICITUD INTERNACIONAL PUBLICADA EN VIRTUD DEL TRATADO DE COOPERACIÓN EN MATERIA DE PATENTES (PCT)

(19) Organización Mundial de la
Propiedad Intelectual
Oficina internacional



(10) Número de Publicación Internacional
WO 2015/177394 A1

(43) Fecha de publicación internacional
26 de noviembre de 2015 (26.11.2015) **WIPO | PCT**

(51) Clasificación Internacional de Patentes:
G01N 33/50 (2006.01) *G01N 33/533* (2006.01)

(74) Mandatario: **FUSTER, Gustavo**; Pº de la Habana, 9-11,
28036 Madrid (ES).

(21) Número de la solicitud internacional:
PCT/ES2015/070402

(81) Estados designados (*a menos que se indique otra cosa, para toda clase de protección nacional admisible*): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(22) Fecha de presentación internacional:
22 de mayo de 2015 (22.05.2015)

(25) Idioma de presentación: español

(26) Idioma de publicación: español

(30) Datos relativos a la prioridad:
P201430755 22 de mayo de 2014 (22.05.2014) ES
P201530122 30 de enero de 2015 (30.01.2015) ES

(84) Estados designados (*a menos que se indique otra cosa, para toda clase de protección regional admisible*): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), euroasiática (AM, AZ, BY, KG, KZ, RU, TJ, TM), europea (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

(71) Solicitantes: **SERVICIO ANDALUZ DE SALUD** [ES/ES]; Avda de la Constitución 18, 41071 Sevilla (ES). **CENTRO DE INVESTIGACIÓN BIOMÉDICA EN RED (CIBER)** [ES/ES]; Monforte de Lemos, 5, Pabellón 11, 28029 Madrid (ES). **FUNDACIÓN PARA LA INVESTIGACIÓN BIOMÉDICA DE CÓRDOBA** [ES/ES]; Avda Menéndez Pidal, s/n, Hospital Universitario Reina Sofía, Edificio Consultas Externas, n-1, 14004 Córdoba (ES).

(72) Inventores: **FERRÍN SÁNCHEZ, Gustavo**; Avda Menéndez Pidal, s/n, Hospital Universitario Reina Sofía, Edificio Consultas Externas, n-1, 14004 Córdoba (ES). **DE LA MATA GARCÍA, Manuel**; Avda Menéndez Pidal, s/n, Hospital Universitario Reina Sofía, Edificio Consultas Externas, n-1, 14004 Córdoba (ES). **RODRÍGUEZ PERÁLVAREZ, Manuel Luis**; Avda Menéndez Pidal, s/n, Hospital Universitario Reina Sofía, Edificio Consultas Externas, n-1, 14004 Córdoba (ES).

Publicada:

- con informe de búsqueda internacional (Art. 21(3))
- con la parte de lista de secuencias de la descripción (Regla 5.2(a))

(54) Title: METHOD FOR OBTAINING USEFUL DATA FOR THE DIAGNOSIS OF HEPATOCELLULAR CARCINOMA IN ALCOHOLIC PATIENTS SUFFERING FROM HEPATITIS C

(54) Título : MÉTODO DE OBTENCIÓN DE DATOS ÚTILES PARA EL DIAGNÓSTICO DE HEPATOCARCINOMA EN PACIENTES ALCOHÓLICOS CON HEPATITIS C

(57) Abstract: The invention relates to a method for obtaining useful data from biological samples, for the diagnosis of hepatocellular carcinoma in alcoholic individuals suffering from hepatitis C. The invention also relates to the corresponding kit and uses.

(57) Resumen: Método de obtención de datos útiles para el diagnóstico de hepatocarcinoma en individuos alcohólicos con hepatitis C a partir de muestras biológicas, kit y usos.



WO 2015/177394 A1

Método de obtención de datos útiles para el diagnóstico de hepatocarcinoma en pacientes alcohólicos con hepatitis C

CAMPO DE LA INVENCION

- 5 La presente invención se encuentra dentro del campo de la biomedicina y la biotecnología, y específicamente se refiere a un método de obtención de datos útiles para diagnosticar el hepatocarcinoma en pacientes alcohólicos que presentan hepatitis C.

ANTECEDENTES DE LA INVENCION

- 10 La infección crónica por el virus de la hepatitis C (VHC) es la mayor causa de enfermedad hepática, cirrosis y carcinoma hepatocelular (HCC) en pacientes alcohólicos (Schiff *et al.*, 2003. *Alcohol Res. Health* 27(3), 232-239). La infección crónica por VHC y el consumo de alcohol persistente son dos factores de riesgo significativos e independientes para el desarrollo de la enfermedad hepática terminal y la progresión del HCC, ambos actúan sinérgicamente para
15 acelerar el daño hepático. Por lo tanto, los pacientes alcohólicos infectados por el VHC tienen una ocurrencia más rápida y frecuente de fibrosis y cirrosis, en comparación con los pacientes no alcohólicos. Algunos de los mecanismos principales que se postulan como responsables de la progresión de la enfermedad son: las altas tasas de apoptosis, la peroxidación de lípidos, la generación de radicales libres y de especies reactivas del oxígeno con la reducción de la
20 capacidad antioxidante del hígado (Siu *et al.*, 2009. *Semin. Liver Dis.* 29(2), 188-199).

- Muchos de los HCC son identificados en etapas muy avanzadas y su pronóstico es muy pobre. Las terapias están dirigidas fundamentalmente para los tumores tempranos, la mayoría de ellos asintomáticos (McKillop *et al.*, 2009. *Semin. Liver Dis.* 29(2), 222-232). Además, los marcadores séricos analizados para la detección del HCC (des-gamma-carboxi-protrombina y alfa
25 fetoproteína) no se utilizan rutinariamente debido a que poseen una sensibilidad y especificidad insuficientes (Lok *et al.*, 2010. *Gastroenterology* 138(2), 493-502; EASL-EORTC, 2012. *J. Hepatol.* 56(4), 908-943). La identificación de biomarcadores para la detección precoz del HCC es esencial para mejorar la supervivencia de los pacientes.

- La electroforesis en gel bidimensional de diferencia de fluorescencia (2-DIGE) para el análisis de
30 muestras clínicas ha sido utilizada para identificar biomarcadores potenciales para la detección temprana y el diagnóstico de cáncer humano, así como para entender los mecanismos implicados en el proceso de carcinogénesis (Na *et al.*, 2012. *Methods Mol. Biol.* 854, 223-237). Determinadas estrategias de reducción han demostrado su efectividad para eliminar la abundancia de proteínas

y la mejora de la detección de proteínas poco abundantes en suero/plasma (Liang *et al.*, 2012. *Methods Mol. Biol.* 854, 181-191).

Con el objetivo principal de evitar las dificultades que se presentan en el diagnóstico temprano del HCC, que se genera mayoritariamente en pacientes con factores de riesgo como la infección crónica por el VHC y el alcoholismo, la presente invención se refiere a un método para asistir al facultativo en la diagnosis del HCC en pacientes con infección por VHC y alcohólicos.

BREVE DESCRIPCIÓN DE LA INVENCION

Un **primer aspecto** de la invención se refiere al uso de al menos uno de los genes *C4A*, *CP*, *PON1*, o cualquiera de sus combinaciones, para la obtención de datos útiles para el diagnóstico del hepatocarcinoma en un individuo con hepatitis C. Preferiblemente, los genes *C4A*, *CP*, *PON1* se usan de manera simultánea.

En otra realización preferida de este aspecto de la invención, el individuo es alcohólico.

Un **segundo aspecto** de la invención se refiere a un método de obtención de datos útiles para el diagnóstico del hepatocarcinoma en un individuo con hepatitis C, de ahora en adelante primer método de la invención, que comprende:

i) obtener una muestra biológica aislada de un individuo

ii) cuantificar la cantidad de producto de expresión de los genes que se seleccionan de la lista: *C4A*, *CP*, *PON1*, o cualquiera de sus combinaciones.

Preferiblemente, en el paso (ii) se determina la cantidad de producto de expresión de todos los genes *C4A*, *CP*, *PON1*. En una realización preferida de este aspecto de la invención, el método además comprende:

iii) hallar un valor de diagnóstico (VD) derivado con los valores obtenidos en ii), mediante la fórmula:

$$VD = (C4A \times CP) / PON1 \text{ (ng/mg de proteína plasmática)}$$

Un **tercer aspecto** de la invención se refiere a un método de diagnóstico del hepatocarcinoma en un individuo con hepatitis C, de ahora en adelante segundo método de la invención, que comprende los pasos (i)-(iii) del primer método de la invención, y además comprende clasificar a los individuos que presentan un $VD > 3,60$ ng/mg, preferiblemente $VD > 3,70$ ng/mg, más preferiblemente $VD > 3,80$ ng/mg, y aún más preferiblemente $VD > 3,87$ ng/mg, en el grupo de individuos con hepatocarcinoma.

En una realización preferida de este aspecto de la invención, la muestra biológica aislada i) es una muestra de sangre, preferiblemente suero.

En otra realización preferida de este aspecto de la invención, el individuo es alcohólico.

Otro aspecto de la invención se refiere a un método de diagnóstico del hepatocarcinoma en un individuo con hepatitis C, que comprende los pasos (i)-(ii) del primer método de la invención, o los pasos (i)-(iii) del segundo método de la invención, y que además comprende: asignar los individuos del paso (i) al grupo de individuos con hepatocarcinoma uninodular cuando presentan un valor de C4A < 2,00 ng/mg, y más preferiblemente C4A < 2,07 ng/mg, o asignar a los individuos del paso (i) al grupo de individuos con hepatocarcinoma multinodular cuando presentan un valor de C4A > 2,00 ng/mg, y más preferiblemente un valor de C4A > 2,07 ng/mg.

En una realización preferida de este aspecto de la invención, la muestra biológica aislada i) es una muestra de sangre, preferiblemente suero.

En otra realización preferida de este aspecto de la invención, el individuo es alcohólico.

En otra realización preferida de este aspecto de la invención, la detección de la cantidad de expresión de cualquiera de los genes C4A, CP y PON1 se realiza mediante un inmunoensayo.

En otra realización preferida, el inmunoensayo es un ensayo inmunoabsorbente ligado a enzimas o ELISA (*Enzyme-Linked ImmunoSorbent Assay*).

Un **cuarto aspecto** de la invención se refiere a un kit o dispositivo que comprende los elementos necesarios para cuantificar la cantidad de expresión de los genes que se seleccionan de entre C4A, CP y PON1.

En una realización preferida de este aspecto de la invención, el kit o dispositivo comprende los anticuerpos que se seleccionan de la lista que consiste en anticuerpos: anti-C4A, anti-CP y anti-PON1.

En otra realización preferida el kit o dispositivo comprende las sondas SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 6 y/o SEQ ID NO: 8.

Un **quinto aspecto** de la invención se refiere al uso del kit o dispositivo para obtener datos útiles para el diagnóstico del hepatocarcinoma en un individuo.

Adicionalmente, los inventores han comprobado como la determinación o cuantificación del producto de expresión del gen FAG en combinación con la determinación o cuantificación del producto de expresión de los genes C4A, CP y PON1 resulta especialmente útil en el diagnóstico del hepatocarcinoma en un individuo con hepatitis C

Por lo tanto, un **sexto aspecto** de la invención se refiere al uso, preferiblemente *in vitro*, de los genes que se seleccionan de la lista *C4A*, *FGA*, *CP*, *PON1*, o de los productos de expresión de dichos genes, o cualquiera de sus combinaciones, para la obtención de datos útiles para el diagnóstico del hepatocarcinoma en un individuo con hepatitis C.

- 5 En una realización preferida de este aspecto de la invención, el uso de los genes *C4A*, *FGA*, *CP*, *PON1*, o de los productos de expresión de dichos genes, es simultáneo.

Un **séptimo aspecto** de la invención se refiere a un método de obtención de datos útiles para el diagnóstico del hepatocarcinoma en un individuo con hepatitis C, de ahora en adelante primer método de la invención, que comprende:

- 10 i) obtener una muestra biológica aislada de un individuo
- ii) cuantificar la cantidad de producto de expresión de los genes que se seleccionan de la lista: *C4A*, *FGA*, *CP*, *PON1*, o cualquiera de sus combinaciones.

En una realización preferida de este aspecto de la invención, se cuantifica simultáneamente la cantidad de producto de expresión de los genes *C4A*, *FGA*, *CP* y *PON1*.

- 15 En otra realización preferida de este aspecto de la invención, el método además comprende:

iii) hallar un valor de diagnóstico (VD) derivado con los valores obtenidos en ii), mediante la fórmula:

$$VD = (C4A \times FGA \times CP) / PON1$$

- 20 Un **octavo aspecto** de la invención se refiere a un método de diagnóstico del hepatocarcinoma en un individuo con hepatitis C, de ahora en adelante método de la invención FGA, que comprende los pasos (i)-(iii) del método del séptimo aspecto de la invención, y además comprende clasificar a los individuos que presentan un $VD > 95$, preferiblemente un $VD > 105$, más preferiblemente un $VD > 116$, y aún más preferiblemente un $VD > 120$, en el grupo de individuos con hepatocarcinoma.

- 25 Un **noveno aspecto** de la invención se refiere a un método de diagnóstico del hepatocarcinoma en un individuo con hepatitis C, que comprende los pasos (i)-(ii) del método del séptimo aspecto de la invención, y que además comprende: asignar los individuos del paso (i) al grupo de individuos con hepatocarcinoma uninodular cuando presentan un valor de producto de expresión del gen *C4A* $< 2,00$ ng/mg, y más preferiblemente un valor de producto de expresión del gen *C4A* $< 2,07$ ng/mg, o asignar a los individuos del paso (i) al grupo de individuos con hepatocarcinoma multinodular cuando presentan un valor de producto de expresión del gen *C4A* $> 2,00$ ng/mg, y más preferiblemente un valor de producto de expresión del gen *C4A* $> 2,07$ ng/mg.
- 30

En una realización preferida de este aspecto de la invención, la muestra biológica aislada i) es una muestra de sangre, preferiblemente plasma sanguíneo.

En otra realización preferida de este aspecto de la invención, el individuo es alcohólico.

En otra realización preferida de este aspecto de la invención, la detección de la cantidad de expresión de cualquiera de los genes *C4A*, *FGA*, *CP* y *PON1* se realiza mediante un inmunoensayo.

En otra realización preferida, el inmunoensayo es un ensayo inmunoabsorbente ligado a enzimas o ELISA (*Enzyme-Linked ImmunoSorbent Assay*).

Un **décimo aspecto** de la invención se refiere a un kit o dispositivo que comprende los elementos necesarios para cuantificar la cantidad de los productos expresión de los genes que se seleccionan de entre *C4A*, *FGA*, *CP* y *PON1*, o cualquiera de sus combinaciones.

Aún más preferiblemente, el kit de este aspecto de la invención comprende anticuerpos o fragmentos de los mismos específicos frente a cualquiera de las secuencias SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 9 o SEQ ID NO: 11 o frente a secuencias aminoacídicas que presenten un grado de identidad con dichas secuencias aminoacídicas de al menos del 85%, típicamente de, al menos del 90%, preferiblemente de, al menos del 95%, más preferiblemente de, al menos del 98%, aún más preferiblemente de, al menos del 99%. En otra realización preferida, el kit de la invención comprende anticuerpos secundarios o controles positivos y/o negativos. El kit además puede incluir, sin ningún tipo de limitación, tampones, soluciones de extracción de proteínas, agentes para prevenir la contaminación, inhibidores de la degradación de las proteínas, etc.

Por otro lado el kit puede incluir todos los soportes y recipientes necesarios para su puesta en marcha y optimización. Preferiblemente, el kit comprende además las instrucciones para llevar a cabo los métodos de la invención.

En otra realización preferida, el kit de este aspecto de la invención es adecuado para amplificar secuencias nucleótidas, que comprende sondas y/o cebadores diseñados a partir de las secuencias SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10 y/o SEQ ID NO: 12 así como opcionalmente todos aquellos elementos necesarios para llevar a cabo un procedimiento PCR.

En una realización preferida, el kit de este aspecto de la invención comprende la asignación de un valor de diagnóstico y clasificación de los individuos, de manera que: si los individuos presentan un VD > 120 se clasifican en el grupo de individuos con hepatocarcinoma.

Un **undécimo aspecto** de la invención se refiere a un soporte sólido, o chip de proteínas, que comprende al menos uno de los anticuerpos o fragmentos de los mismos específicos frente a cualquiera de las secuencias SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 9 o SEQ ID NO: 11 o frente a secuencias aminoacídicas que presenten un grado de identidad con dichas secuencias aminoacídicas de al menos del 85%, típicamente de, al menos del 90%, preferiblemente de, al menos del 95%, más preferiblemente de, al menos del 98%, aún más preferiblemente de, al menos del 99%, o cualquiera de sus combinaciones o cualquiera de sus combinaciones, para llevar a cabo el cualquiera de los métodos de la invención.

Un **duodécimo aspecto** de la invención se refiere a un soporte sólido, o chip de ADN, que comprende oligonucleótidos o microarreglos de canal único diseñados a partir de una secuencia conocida o un ARNm de al menos uno de los genes *C4A*, *FGA*, *CP* y *PON1*. Preferiblemente, comprende los oligonucleótidos capaces de detectar el ARNm de todos los genes *C4A*, *FGA*, *CP* y *PON1*.

Un **décimo tercer aspecto** de la invención se refiere al uso del kit o dispositivo de la invención, en la obtención de datos útiles para el diagnóstico del hepatocarcinoma en un individuo con hepatitis C, preferiblemente alcohólico.

Un **décimo cuarto aspecto** de la invención se refiere a un medio de almacenamiento legible por un ordenador que comprende instrucciones de programa capaces de hacer que un ordenador lleve a cabo los pasos de cualquiera de los métodos de la invención.

DESCRIPCIÓN DE LAS FIGURAS

Figura 1. Imágenes del gel. (a) Superposición maestra del gel que contiene proteínas etiquetadas Cy2, Cy3 y Cy5 de las muestras de plasma. (b) SYPRO® Ruby (BioRad) teñidas de imagen del gel preparativa. Están marcadas las manchas de proteínas identificadas expresadas diferencialmente.

Figura 2. Confirmación de los resultados DIGE por western-blot. Detección y cuantificación de (A) carboxipeptidasa-N (CPN), (B) paraoxonasa sérica/arilesterasa 1 (PON1), y (C) ceruloplasmina (CP) en muestras de plasma humano analizados previamente por DIGE. La densidad de banda de proteína se calculó con QuantityOne 4.4.0 software (BioRad). Los resultados se muestran como unidades arbitrarias (AU), (D) proteínas transferidas a la membrana de nitrocelulosa se detectaron mediante tinción de Ponceau S como control de carga western-blot.

Figura 3. Validación de biomarcadores candidatos seleccionados para el desarrollo de HCC en una población de pacientes mayor. La cuantificación de la concentración plasmática de (A)

CPN (pg/mg de proteína de plasma), PON1 y CP ($\mu\text{g/mg}$ de proteína de plasma), y C4A (pg/mg de proteína de plasma) se realizó por test de ELISA. El tamaño de la muestra fue de 24 en el grupo de control y 22 en el grupo de tumores. La significación estadística se determinó mediante la prueba *T* de Student ($p < 0,05$). Los datos se expresan como media $\pm$ desviación estándar; curvas (C) ROC muestran que la mejor capacidad predictiva para identificar a los pacientes con HCC pertenece a la combinación de C4A, PC y PON1.

Figura 4. Validación de biomarcadores candidatos seleccionados para el desarrollo de HCC en una población de pacientes mayor. (A) La cuantificación de la concentración plasmática de CPN (pg/mg de proteína de plasma), CP, IgM, PON1 ($\mu\text{g/mg}$ de proteína de plasma), C4A (ng/mg de proteína de plasma) y FGA ($\mu\text{g/mg}$ de proteína de plasma) se realizó por test de ELISA. El tamaño de la muestra fue de 24 en el grupo de control y 22 en el grupo de tumores. La significación estadística se determinó mediante la prueba *T* de Student ($p < 0,05$). Los datos se expresan como media $\pm$ error estándar; (B) Las curvas ROC muestran que la mejor capacidad predictiva para identificar a los pacientes con HCC pertenece a la combinación de C4A, FGA, CP y PON1.

DESCRIPCIÓN DETALLADA DE LA INVENCION

Los autores de la presente invención han analizado la expresión de los genes C4A, FGA, CP y PON1 en individuos alcohólicos con hepatitis C y han generado un método para la obtención de datos útiles para el diagnóstico del hepatocarcinoma

La presente invención se refiere a un método y a un kit o dispositivo para su uso, preferiblemente *in vitro*, en el diagnósticos del hepatocarcinoma en individuos alcohólicos con hepatitis C.

Por tanto, un aspecto que denominaremos **A** de la invención se refiere al uso de al menos uno de los genes que se seleccionan de la lista C4A, FGA, CP, PON1, o de los productos de expresión de dichos genes, o cualquiera de sus combinaciones, para la obtención de datos útiles para el diagnóstico del hepatocarcinoma en un individuo con hepatitis C.

En una realización preferida de este aspecto de la invención, el uso de los genes C4A, FGA, CP, PON1, o de los productos de expresión de dichos genes, es secuencial o simultáneo. El producto de expresión de cada uno de dichos genes se puede determinar de forma simultánea o secuencial en cualquier orden.

En una realización preferida, por producto de expresión del gen C4A, se entiende la secuencia aminoácidica que presente un grado de identidad con la secuencia aminoácidica SEQ ID NO: 1 o SEQ ID NO: 2 de, al menos del 85%, típicamente de, al menos del 90%, preferiblemente de, al

menos del 95%, más preferiblemente de, al menos del 98%, aún más preferiblemente de, al menos del 99% o una secuencia nucleotídica que codifique para cualquiera de dichas secuencia aminoácidas.

- 5 En una realización preferida, por producto de expresión del gen FGA, se entiende las secuencia aminoácida que presente un grado de identidad con la secuencia aminoácida SEQ ID NO: 5 o SEQ ID NO: 6 de, al menos del 85%, típicamente de, al menos del 90%, preferiblemente de, al menos del 95%, más preferiblemente de, al menos del 98%, aún más preferiblemente de, al menos del 99% o una secuencia nucleotídica que codifique para cualquiera de dichas secuencia aminoácidas.
- 10 En una realización preferida, por producto de expresión del gen CP, se entiende las secuencia aminoácida que presente un grado de identidad con la secuencia aminoácida SEQ ID NO: 9 de, al menos del 85%, típicamente de, al menos del 90%, preferiblemente de, al menos del 95%, más preferiblemente de, al menos del 98%, aún más preferiblemente de, al menos del 99% o una secuencia nucleotídica que codifique para cualquiera de dichas secuencia aminoácidas.
- 15 En una realización preferida, por producto de expresión del gen PON1, se entiende las secuencia aminoácida que presente un grado de identidad con la secuencia aminoácida SEQ ID NO: 11 de, al menos del 85%, típicamente de, al menos del 90%, preferiblemente de, al menos del 95%, más preferiblemente de, al menos del 98%, aún más preferiblemente de, al menos del 99% o una secuencia nucleotídica que codifique para cualquiera de dichas secuencia aminoácidas.
- 20 En otra realización preferida de este aspecto de la invención, el individuo es alcohólico.

El término "alcohólico", tal y como se utiliza en la descripción, se refiere a individuos con un consumo diario regular de etanol superior a 40g en hombres y a 20g en mujeres. El consumo de riesgo es un nivel o patrón de consumo de alcohol que puede causar daños en la salud si el hábito del consumo persiste, y es descrito por la Organización Mundial de la Salud (OMS) como

25 consumo medio regular de 20 a 40g de alcohol diarios en mujeres, y de 40 a 60g diarios en hombres. La dependencia del alcohol es un conjunto de fenómenos conductuales, cognitivos y fisiológicos en los cuales el uso del alcohol se transforma en prioritario para el individuo, en contraposición con otras conductas que en algún momento tuvieron mayor valor para él. El consumo lleva a un estado de intoxicación y el alcoholismo lleva asociado un conjunto de

30 alteraciones digestivas (gastritis, cirrosis que es causa de insuficiencia hepática, hemorragias o encefalopatía de pronóstico grave), favorece el desarrollo de cáncer de esófago, produce atrofia cerebral, neuritis de nervios periféricos, neuritis óptica y, en casos avanzados, conduce a la pérdida de reflejos, disminución de la sensibilidad, trastornos tróficos incluso parálisis y ceguera total.

Un **aspecto B** de la invención se refiere a un método de obtención de datos útiles para el diagnóstico del hepatocarcinoma en un individuo con hepatitis C, de ahora en adelante primer método de la invención, que comprende:

i) obtener una muestra biológica aislada de un individuo;

5 ii) cuantificar la cantidad de producto de expresión de los genes que se seleccionan de la lista: *C4A*, *CP*, *PON1*, o cualquiera de sus combinaciones a partir de la muestra biológica del individuo del paso i).

En una realización preferida de este aspecto de la invención, se cuantifica simultáneamente la cantidad de producto de expresión de los genes *C4A*, *CP* y *PON1*.

10 En otra realización preferida de este aspecto de la invención, el método además comprende:

iii) hallar un valor de diagnóstico (VD) derivado con los valores obtenidos en ii), mediante la fórmula:

$$VD = (C4A \times CP) / PON1 \text{ (ng/mg de proteína plasmática)}$$

15 En una realización preferida de este aspecto de la invención, la muestra biológica aislada i) es una muestra de sangre, y más preferiblemente es suero.

En otra realización preferida de este aspecto de la invención, el individuo es alcohólico.

20 Un **aspecto C** de la invención se refiere a un método de diagnóstico del hepatocarcinoma en un individuo con hepatitis C, de ahora en adelante segundo método de la invención, que comprende los pasos (i)-(iii) del primer método de la invención, y además comprende clasificar a los individuos que presentan un $VD > 3,87$ ng/mg en el grupo de individuos con hepatocarcinoma.

En una realización preferida de este aspecto de la invención, la muestra biológica aislada i) es una muestra de sangre, preferiblemente suero.

En otra realización preferida de este aspecto de la invención, el individuo es alcohólico.

25 Otro aspecto de la invención se refiere a un método de diagnóstico del hepatocarcinoma en un individuo con hepatitis C, que comprende los pasos (i)-(ii) del primer método de la invención o los pasos (i)-(iii) del segundo método de la invención, y que además comprende: asignar los individuos del paso (i) al grupo de individuos con hepatocarcinoma uninodular cuando presentan un valor de $C4A < 2,07$ ng/mg, o asignar a los individuos del paso (i) al grupo de individuos con hepatocarcinoma multinodular cuando presentan un valor de $C4A > 2,07$ ng/mg.

30 En una realización preferida de este aspecto de la invención, la muestra biológica aislada i) es una muestra de sangre, preferiblemente suero.

En otra realización preferida de este aspecto de la invención, el individuo es alcohólico.

Un aspecto D de la invención se refiere a un método de obtención de datos útiles para el diagnóstico del hepatocarcinoma en un individuo con hepatitis C, de ahora en adelante tercer método de la invención, que comprende:

- 5 i) obtener una muestra biológica aislada de un individuo
- ii) cuantificar la cantidad de producto de expresión de los genes que se seleccionan de la lista: C4A, FGA, CP, PON1, o cualquiera de sus combinaciones.

En una realización preferida de este aspecto de la invención, se cuantifica simultáneamente la cantidad de producto de expresión de los genes C4A, FGA, CP y PON1.

- 10 En otra realización preferida de este aspecto de la invención, el método además comprende:

 iii) hallar un valor de diagnóstico (VD) derivado con los valores obtenidos en ii), mediante la fórmula:

$$VD = (C4A \times FGA \times CP) / PON1$$

- 15 En una realización preferida de este aspecto de la invención, la muestra biológica aislada i) es una muestra de sangre, y más preferiblemente es plasma, preferiblemente sanguíneo.

En otra realización preferida de este aspecto de la invención, el individuo es alcohólico.

- 20 **Un aspecto E** de la invención se refiere a un método de diagnóstico del hepatocarcinoma en un individuo con hepatitis C, de ahora en adelante cuarto método de la invención, que comprende los pasos (i)-(iii) del tercer método de la invención, y además comprende clasificar a los individuos que presentan un $VD > 95$, preferiblemente un $VD > 105$, más preferiblemente un $VD > 116$, y aún más preferiblemente un $VD > 120$, en el grupo de individuos con hepatocarcinoma.

En una realización preferida de este aspecto de la invención, la cantidad del producto de expresión de los genes se mide: en ng/mg para el gen C4A y en µg/mg para los genes FGA, CP, PON1.

- 25 En una realización preferida de este aspecto de la invención, la muestra biológica aislada i) es una muestra de sangre, preferiblemente plasma sanguíneo.

En otra realización preferida de este aspecto de la invención, el individuo es alcohólico.

- 30 **Un aspecto F** de la invención se refiere a un método de diagnóstico del hepatocarcinoma en un individuo con hepatitis C, que comprende los pasos (i)-(ii) del tercer método de la invención, y que además comprende: asignar los individuos del paso (i) al grupo de individuos con

hepatocarcinoma uninodular cuando presentan un valor de producto de expresión del gen C4A < 2,00 ng/mg, y más preferiblemente un valor de producto de expresión del gen C4A < 2,07 ng/mg, o asignar a los individuos del paso (i) al grupo de individuos con hepatocarcinoma multinodular cuando presentan un valor de producto de expresión del gen C4A > 2,00 ng/mg, y más preferiblemente un valor de producto de expresión del gen C4A > 2,07 ng/mg.

En una realización preferida de este aspecto de la invención, el método comprende los pasos (i)-(iii) del tercer método de la invención.

En una realización preferida de este aspecto de la invención, la muestra biológica aislada i) es una muestra de sangre, preferiblemente plasma sanguíneo.

10 En otra realización preferida de este aspecto de la invención, el individuo es alcohólico.

Los pasos (ii) y/o (iii) de cualquiera de los métodos descritos anteriormente pueden ser total o parcialmente automatizados, por ejemplo, por medio de un equipo robótico sensor para la detección de la cantidad en el paso (ii) o la comparación computarizada en el paso (iii).

15 Una “muestra biológica aislada” incluye, pero sin limitarnos a, células, tejidos y/o fluidos biológicos de un organismo, obtenidos mediante cualquier método conocido por un experto en la materia. Preferiblemente, la muestra biológica aislada es una muestra de sangre, y más preferiblemente es plasma sanguíneo.

20 En esta memoria se entiende por “plasma sanguíneo” a la fracción líquida y acelular de la sangre obtenida al dejarla desprovista de células. Normalmente, está compuesto por un 90 % de agua, un 7 % de proteínas, y el 3 % restante por grasa, glucosa, vitaminas, hormonas, oxígeno, gas carbónico y nitrógeno, además de productos de desecho del metabolismo como el ácido úrico. A estos se les pueden añadir otros compuestos como las sales y la urea. El plasma sanguíneo representa aproximadamente el 55% del volumen sanguíneo total, mientras que el 45 % restante corresponde a los elementos formes. Las principales proteínas plasmáticas se clasifican en:

25 albúminas, globulinas y fibrinógeno.

El término “individuo”, tal y como se utiliza en la descripción, se refiere a animales, preferiblemente mamíferos, y más preferiblemente, humanos. El término “individuo” no pretende ser limitativo en ningún aspecto, pudiendo ser éste de cualquier edad, sexo y condición física.

30 La detección la cantidad de producto de expresión de C4A, FGA, CP y PON1 puede realizarse por cualquier medio conocido en el estado de la técnica. Los autores de la presente invención han demostrado que la detección de la cantidad o la concentración de anticuerpos frente a estas proteínas de manera semi-cuantitativa o cuantitativa permiten la obtención de datos útiles para el diagnóstico del hepatocarcinoma en individuos con hepatitis C.

Los niveles de expresión de los genes van a dar un determinado perfil de expresión génica. El término “nivel de expresión”, también denominado “cantidad producto génico” o “cantidad de producto de expresión” se refiere al material bioquímico, ya sea ARN o proteína, resultado de la expresión de un gen. Algunas veces se usa una medida de la cantidad de producto génico para inferir qué tan activo es un gen. Se entiende por “perfil de expresión génica” el perfil génico obtenido tras la cuantificación del ARNm y/o de proteína producida por los genes de interés o biomarcadores, es decir, por los genes C4A, FGA, CP y PON1.

La medida de la cantidad o la concentración de producto de expresión de estos genes, preferiblemente de manera semi-cuantitativa o cuantitativa, puede ser llevada a cabo de manera directa o indirecta. La medida directa se refiere a la medida de la cantidad o la concentración del producto de expresión de los genes, basada en una señal que se obtiene directamente de los transcritos de dichos genes, o de las proteínas, y que está correlacionada directamente con el número de moléculas de ARN o de proteínas producidas por los genes. Dicha señal (a la que también podemos referirnos como señal de intensidad) puede obtenerse, por ejemplo, midiendo un valor de intensidad de una propiedad química o física de dichos productos. La medida indirecta incluye la medida obtenida de un componente secundario o un sistema de medida biológica (por ejemplo la medida de respuestas celulares, ligandos, “etiqueta” o productos de reacción enzimática).

El término “cantidad”, tal y como se utiliza en la descripción, se refiere pero no se limita, a la cantidad absoluta o relativa de los productos de expresión de los genes o a la cantidad de las proteínas, así como a cualquier otro valor o parámetro relacionado con los mismos o que pueda derivarse de éstos. Dichos valores o parámetros comprenden valores de intensidad de la señal obtenidos a partir de cualquiera de las propiedades físicas o químicas de dichos productos de expresión obtenidos mediante medida directa. Adicionalmente, dichos valores o parámetros incluyen todos aquellos obtenidos mediante medida indirecta, por ejemplo, cualquiera de los sistemas de medida descritos en otra parte del presente documento.

El término “comparación”, tal y como se utiliza en la descripción, se refiere pero no se limita, a la comparación de la cantidad de los productos de expresión de los genes o de la cantidad de anticuerpos frente a C4A, FGA, CP y PON1 y de la muestra biológica a analizar, también llamada muestra biológica problema, con una cantidad de los productos de expresión de los genes o con una cantidad de anticuerpos frente a C4A, FGA, CP y PON1 de una o varias muestras de referencia deseable. La muestra de referencia puede ser analizada, por ejemplo, simultánea o consecutivamente, junto con la muestra biológica problema. El cálculo descrito en el apartado (iii) del método de la presente invención puede ser realizado manualmente o asistido por ordenador.

En la siguiente tabla se describen los genes usados en la invención según la base de datos PubMed del NCBI (National Center of Biotechnology Information):

Nombre del Gen	Descripción	Gen ID	Secuencias aminoacídicas	Secuencias nucleotídicas
<i>C4A</i>	Componente del complemento 4a	720	SEQ ID NO: 1 o SEQ ID NO: 2	SEQ ID NO: 3 o SEQ ID NO: 4
<i>FGA</i>	Fibrinógeno alfa	2243	SEQ ID NO: 5 o SEQ ID NO: 6	SEQ ID NO: 7 o SEQ ID NO: 8
<i>CP</i>	Ceruloplasmina (ferredoxina)	1356	SEQ ID NO: 9	SEQ ID NO: 10
<i>PON1</i>	Paraoxonasa 1	5444	SEQ ID NO: 11	SEQ ID NO: 12

A continuación se describen las características de los genes y de los productos de expresión o proteínas que son objeto del estudio:

C4A: este gen codifica para la forma ácida del factor de complemento 4, parte de la ruta de activación clásica. La proteína se expresa como un único precursor de cadena simple que se escinde proteolíticamente en un trímero de cadenas alfa, beta y gamma antes de la secreción. El trímero proporciona una superficie de interacción entre el complejo antígeno-anticuerpo y otros componentes del complemento. La cadena alfa se puede escindir para liberar anafilatoxina C4, que actúa como mediador de la inflamación local. La deficiencia de esta proteína se asocia con lupus eritematoso sistémico y diabetes mellitus tipo I (Lhotta et al., 1996. Diabetes Care 19(1), 53-55). Este gen se localiza en la región del complejo mayor de histocompatibilidad (MHC) clase III del cromosoma 6. Existen diferentes haplotipos de este grupo de genes, de tal manera que los individuos pueden tener 1, 2 ó 3 copias de este gen. Para este gen se han encontrado dos variantes de transcripción que codifican diferentes isoformas (PubMed Gene ID: 720).

En el contexto de la presente invención, C4A se define también por una secuencia de nucleótidos o polinucleótidos, que constituye la secuencia codificante de la proteína recogida en la SEQ ID NO: 1 o SEQ ID NO: 2, y que comprendería diversas variantes procedentes de:

- a) moléculas de ácido nucleico que codifican un polipéptido que comprende la secuencia aminoacídica de la SEQ ID NO: 1 o SEQ ID NO: 2 ,

b) moléculas de ácido nucleico cuya cadena complementaria híbrida con la secuencia polinucleotídica de a),

c) moléculas de ácido nucleico cuya secuencia difiere de a) y/o b) debido a la degeneración del código genético,

5 d) moléculas de ácido nucleico que codifican un polipéptido que comprende la secuencia aminoacídica con una identidad de al menos un 80%, un 90%, un 95%, un 98% o un 99% con la SEQ ID NO: 1 o SEQ ID NO: 2, y en las que el polipéptido codificado por dichos ácidos nucleicos posee la actividad y las características estructurales de la proteína C4A. Entre dichas moléculas de ácido nucleico se encuentra la recogida en la secuencia SEQ ID
10 NO: 3 o SEQ ID NO: 4.

FGA: la proteína codificada por este gen es el componente alfa del fibrinógeno, una glicoproteína transmitida por la sangre que está compuesta de tres pares de cadenas polipeptídicas no idénticas. Después de una lesión vascular, el fibrinógeno se escinde por la trombina para formar fibrina, que es el componente más abundante de coágulos de sangre. Además, varios productos
15 de escisión de fibrinógeno y fibrina regulan la adhesión celular y la difusión, la vasoconstricción y actividades quimiotácticas, y son mitógenos para varios tipos de células. Las mutaciones en este gen conducen a varios trastornos, incluyendo disfibrinogenemia, hipofibrinogenemia, afibrinogenemia y amiloidosis renal. Del splicing alternativo resultan dos isoformas que varían en el extremo carboxi-terminal (PubMed Gene ID: 2243).

20 En el contexto de la presente invención, FGA se define también por una secuencia de nucleótidos o polinucleótidos, que constituye la secuencia codificante de la proteína recogida en la SEQ ID NO: 5 o SEQ ID NO: 6, y que comprendería diversas variantes procedentes de:

a) moléculas de ácido nucleico que codifican un polipéptido que comprende la secuencia aminoacídica de la SEQ ID NO: 5 o SEQ ID NO: 6,

25 b) moléculas de ácido nucleico cuya cadena complementaria híbrida con la secuencia polinucleotídica de a),

c) moléculas de ácido nucleico cuya secuencia difiere de a) y/o b) debido a la degeneración del código genético,

30 d) moléculas de ácido nucleico que codifican un polipéptido que comprende la secuencia aminoacídica con una identidad de al menos un 80%, un 90%, un 95%, un 98% o un 99% con la SEQ ID NO: 5 o SEQ ID NO: 6, y en las que el polipéptido codificado por dichos ácidos nucleicos posee la actividad y las características estructurales de la proteína FGA.

Entre dichas moléculas de ácido nucleico se encuentra la recogida en la secuencia SEQ ID NO: 7 o SEQ ID NO: 8.

CP (o también conocida como ferroxidasa): la proteína codificada por este gen es una metaloproteína que une la mayor parte del cobre en el plasma y está implicada en la peroxidación de la transferrina Fe(II) a transferrina Fe(III). Fue descrita por primera vez en 1948 y es la principal proteína transportadora de cobre en la sangre. Las mutaciones en esta proteína causan aceruloplasminemia, caracterizada por la sobrecarga de cobre en el cerebro, páncreas y retina.

En el contexto de la presente invención, CP se define también por una secuencia de nucleótidos o polinucleótidos, que constituye la secuencia codificante de la proteína recogida en la SEQ ID NO: 9, y que comprendería diversas variantes procedentes de:

a) moléculas de ácido nucleico que codifican un polipéptido que comprende la secuencia aminoacídica de la SEQ ID NO: 9,

b) moléculas de ácido nucleico cuya cadena complementaria híbrida con la secuencia polinucleotídica de a),

c) moléculas de ácido nucleico cuya secuencia difiere de a) y/o b) debido a la degeneración del código genético,

d) moléculas de ácido nucleico que codifican un polipéptido que comprende la secuencia aminoacídica con una identidad de al menos un 80%, un 90%, un 95%, un 98% o un 99% con la SEQ ID NO: 9, y en las que el polipéptido codificado por dichos ácidos nucleicos posee la actividad y las características estructurales de la proteína CP. Entre dichas moléculas de ácido nucleico se encuentra la recogida en la secuencia SEQ ID NO: 10.

PON1: La enzima codificada por este gen es una arilesterasa que hidroliza paroxon para producir p-nitrofenol. Los polimorfismos en este gen son un factor de riesgo en enfermedades de arterias coronarias. En humanos, este gen se encuentra en un cluster de genes relacionados con la paraxonasa en la localización cromosómica 7q21.3. La enzima PON1 (arildialquilfosfatasa, E.C.3.1.8.1), descrita en 1946 por A. Mazur (Mazur A. 1946. J. Biol. Chem. 164, 271-89) es dependiente de calcio con actividades de esterasa y lactonasa. Aún sin conocerse su papel fisiológico, se la describió por su capacidad de hidrolizar derivados de compuestos organofosfato; los mismos que inhiben la acción de la acetil colinesterasa, causando síndrome colinérgico y polineuropatía crónica (Lotti M. 1991. Crit. Rev. Toxicol. 21,465-87). Los sustratos de paraoxonasa-1 (PON1) son tri-ésteres del ácido fosfórico como paraoxón y diazoxón, metabolitos de los insecticidas altamente tóxicos para tión y diazinón, respectivamente. Esta enzima hidroliza también los agentes nerviosos sarín y somán y ésteres aromáticos como fenilacetato y naftilacetato. Se ha sugerido que la enzima podría ser capaz de impedir o limitar la oxidación de

las LDL (Mackness et al., 1991 FEBS Lett 286,152-4). Ese fue el primer informe que le adjudicó a la PON1 capacidad antioxidante, tornándose en un punto nodal para entender la importancia clínica de esta enzima, dada la relación entre el estrés oxidativo, la oxidación del LDL y la arteriosclerosis.

5 En el contexto de la presente invención, PON1 se define también por una secuencia de nucleótidos o polinucleótidos, que constituye la secuencia codificante de la proteína recogida en la SEQ ID NO: 11, y que comprendería diversas variantes procedentes de:

a) moléculas de ácido nucleico que codifican un polipéptido que comprende la secuencia aminoacídica de la SEQ ID NO: 11,

10 b) moléculas de ácido nucleico cuya cadena complementaria híbrida con la secuencia polinucleotídica de a),

c) moléculas de ácido nucleico cuya secuencia difiere de a) y/o b) debido a la degeneración del código genético,

15 d) moléculas de ácido nucleico que codifican un polipéptido que comprende la secuencia aminoacídica con una identidad de al menos un 80%, un 90%, un 95%, un 98% o un 99% con la SEQ ID NO: 9, y en las que el polipéptido codificado por dichos ácidos nucleicos posee la actividad y las características estructurales de la proteína PON1. Entre dichas moléculas de ácido nucleico se encuentra la recogida en la secuencia SEQ ID NO: 12.

A continuación se detallan cada una de las secuencias mencionadas.

20 SEQ ID NO:1:

MRLLWGLIWASSFFTLSQLKPRLLLFSPSVVHLGVPLSVGVQLQDVPRGQVVKGSVFLRNPSRNN
VPCSPKVDFTLSSERDFALLSLQVPLKDAKSCGLHQLLRGPEVQLVAHSPWLKDSLSRTTNIQGIN
LLFSSRRGHLFLQTDQPIYNPGQVRVRYRVFALDQKMRPSTDITVMVENSHGLRVRKKEVYMPSSI
FQDDFVIPDISEPGTWKISARFSDGLESNSSTQFEVKKYVLPNFEVKITPGKPYILTVPGHLDQMQL
25 DIQARYIYGKPVQGVAYVRFGLLDEDGKKTFFRGLESQTKLVNGQSHISLSKAEFQDALEKLNMGIT
DLQGLRLYVAAAIIESPGGEMEEAELTSWYFVSSPFSLDLSKTKRHLVPGAPFLLQALVREMSGSPA
SGIPVKVSATVSSPGSVPEVQDIQQNTDGGSGQVSIIIPQTISELQLSVSAGSPHPAIARLTVAAPPS
GGPGFLSIERPDSRPPRVGDTLNLNLRAVGSGATFSHYYYMILSRGQIVFMNREPKRTLTSVSVFV
DHHLAPSFYFVAFYYHGDHPVANSRLVDVQAGACEGKLELSVDGAKQYRNGESVKLHLETDSLAL
30 VALGALDTALYAAGSKSHKPLNMGKVFEAMNSYDLGCGPGGGDSALQVFQAAGLAFSDGDQWTL
SRKRLSCPKEKTTRKKRNVNFQKAINKLGQYASPTAKRCCQDGVTRLPMMSRCEQRAARVQQP
DCREPFLSCCQFAESLRKKS RDKGQAGLQRALEILQEEDLIDEDDIPVRSFFPENWLWRVETVDRF
QILTLWLPDSLTTWEIHGLSLSKTKGLCVATPVQLRVFREFHLHLRLPMSVRRFEQLELRPVLYNYL

DKNLTVSVHVSPVEGLCLAGGGGLAQQVLVPAGSARPVAFSVVPTAAAASLKVVARGSFEFPVG
DAVSKVLQIEKEGAIHREELVYELNPLDHRGRTLEIPGNSDPNMIPDGFNSYVRVTASDPLDTLGS
EGALSPGGVASLLRLPRGCGEQMTMIYLAPTAAASRYLDKTEQWSTLPPETKDHAVDLIQGYMRIQ
QFRKADGSYAAWLSRDSSTWLTAFVLKVLSLAQEQVGGSPPEKLQETSNWLLSQQQADGSFQDPC
5 PVLDRSMQGGGLVGNDDETVALTAFVTIALHHGLAVFQDEGAEPKQRVEASISKANSFLGEKASAGL
LGAHAAAITAYALTTLKAPVDLLGVAHNNLMAMAQETGDNLYWGSVTGSQSNAVSPTPAPRNPSPD
PMPQAPALWIETTAYALLHLLHEGKAEMADQASAWLTRQGSFQGGFRSTQDTVIALDALSAYWIA
SHTTEERGLNVTLSSTGRNGFKSHALQLNNRQIRGLEEELQFSLGSKINVKVGGNSKGTLKVLRTY
NVLDMKNNTTCQDLQIEVTVKGHVETMEANEDYEDYEDYDELPAKDDPDAPLQPVTPPLQLFEGRRN
10 RRRREAPKVVEEQESRVHYTVCIWRNGKVGLSGMAIADVTLTSGFHALRADLEKLTSLSDRYVSH
FETEGPHVLLYFDSVPTSRECVGFQAVQEVVGLVQPASATLYDYNNPERRCSVFYGAPSKSRLLA
TLCSAEVCQCAEGKCPRQRRLERGLQDEDDGYRMKFACYYPVEYGFQVKVLREDSRAAFRLFE
TKITQVLHFTKDVKAAANQMRNFLVRASCRLRLEPGKEYLIMGLDGATYDLEGHPQYLLDSNSWIE
EMPSERLCRSTRQRAACAQLNDFLQEYGTQGCQV

15 SEQ ID NO:2:

MRLLWGLIWASSFFTLSLQKPRLLLFSPSVVHLGVPLSVGVQLQDVPRGQVVKGSVFLRNPSRNN
VPCSPKVDFTLSSERDFALLSLQVPLKDAKSCGLHQLLRGPEVQLVAHSPWLKDSLRTTNIQGIN
LLFSSRRGHLFLQTDQPIYNPGQVRVRYRVFALDQKMRPSTDITVMVENSHGLRVRKKEVYMPSSI
FQDDFVIPDISEPGTWKISARFSDGLESNSSTQFEVKKYVLPNFEVKITPGKPYILTPVGHLEDEMQL
20 DIQARYIYGKPVQGVAYVRFGLLDEDGKKTFFRGLESQTKLVNGQSHISLSKAEFQDALEKLNMGIT
DLQGLRLYVAAAIIESPGGEMEEAELTSWYFVSSPFLDLSKTKRHLVPGAPFLLQALVREMSGSPA
SGIPVKVSATVSSPGSVPEVQDIQQNTDGGGQVSIPIIPQTISELQLSVSAGSPHPAIARLTVAAPPS
GGPGFLSIERPDSRPPRVGDTLNLNLRAVGSGATFSHYYYMILSRGQIVFMNREPRTLTSSVSFV
DHHLAPSFYFVAFYHGDHPVANSRLVDVQAGACEGKLELSVDGAKQYRNGESVKLHLETDSLAL
25 VALGALDTALYAAGSKSHKPLNMGKVFEAMNSYDLGCGPGGGDSALQVFQAAGLAFSDGDQWTL
SRKRLSCPKEKTTRKKRNVNFKKAINELGQYASPTAKRCCQDGVTRLPMMSRCEQRAARVQQP
DCREPFLSCCQFAESLRKKSRDKGQAGLQRALEILQEEDLIDEDDIPVRSFFPENWLWRVETVDRF
QILTLWLPSLTTWEIHGLSLSKTKGLCVATPVQLRVFREFHLHLRLPMSVRRFEQLELRPVLYNYL
DKNLTVSVHVSPVEGLCLAGGGGLAQQVLVPAGSARPVAFSVVPTAAAASLKVVARGSFEFPVG
30 DAVSKVLQIEKEGAIHREELVYELNPLDHRGRTLEIPGNSDPNMIPDGFNSYVRVTASDPLDTLGS
EGALSPGGVASLLRLPRGCGEQMTMIYLAPTAAASRYLDKTEQWSTLPPETKDHAVDLIQGYMRIQ
QFRKADGSYAAWLSRDSSTWLTAFVLKVLSLAQEQVGGSPPEKLQETSNWLLSQQQADGSFQDPC
PVLDRSMQGGGLVGNDDETVALTAFVTIALHHGLAVFQDEGAEPKQRVEASISKANSFLGEKASAGL
LGAHAAAITAYALTTLKAPVDLLGVAHNNLMAMAQETGDNLYWGSVTGSQSNAVSPTPAPRNPSPD
35 PMPQAPALWIETTAYALLHLLHEGKAEMADQASAWLTRQGSFQGGFRSTQDTVIALDALSAYWIA
SHTTEERGLNVTLSSTGRNGFKSHALQLNNRQIRGLEEELQFSLGSKINVKVGGNSKGTLKVLRTY

GTATGCTGCAGGCAGCAAGTCCCACAAGCCCCTCAACATGGGCAAGGTCTTTGAAGCTATGAA
CAGCTATGACCTCGGCTGTGGTCCTGGGGGTGGGGACAGTGCCCTTCAGGTGTTCCAGGCA
GCGGGCCTGGCCTTTTCTGATGGAGACCAGTGACCTTATCCAGAAAGAGACTAAGCTGTCC
CAAGGAGAAGACAACCCGAAAAAGAGAAACGTGAACTTCCAAAAGGCGATTAATGAGAAATT
5 GGGTCAGTATGCTTCCCCGACAGCCAAGCGCTGCTGCCAGGATGGGGTGACACGTCTGCCCA
TGATGCGTTCTGCGAGCAGCGGGCAGCCCGCGTGACAGAGCCGACTGCCGGGAGCCCT
TCCTGTCCTGCTGCCAATTTGCTGAGAGTCTGCGCAAGAAGAGCAGGGACAAGGGCCAGGC
GGGCCTCCAACGAGCCCTGGAGATCCTGCAGGAGGAGGACCTGATTGATGAGGATGACATTC
CCGTGCGCAGCTTCTTCCCAGAGAACTGGCTCTGGAGAGTGGAAACAGTGGACCGCTTTCAA
10 ATATTGACACTGTGGCTCCCCGACTCTCTGACCACGTGGGAGATCCATGGCCTGAGCCTGTCC
AAAACCAAAGGCCTATGTGTGGCCACCCCAGTCCAGCTCCGGGTGTTCCGCGAGTTCACCT
GCACCTCCGCCTGCCCATGTCTGTCCGCCGCTTTGAGCAGCTGGAGCTGCGGCCTGTCCTCT
ATAACTACCTGGATAAAACCTGACTGTGAGCGTCCACGTGTCCCCAGTGGAGGGGCTGTGC
CTGGCTGGGGGCGGAGGGCTGGCCCAGCAGGTGCTGGTGCCTGCGGGCTCTGCCCGGCCT
15 GTTGCTTCTCTGTGGTGCCACGGCAGCCGCCGCTGTGTCTCTGAAGGTGGTGGCTCGAG
GGTCCTTCGAATTCCCTGTGGGAGATGCGGTGTCCAAGGTTCTGCAGATTGAGAAGGAAGGG
GCCATCCATAGAGAGGAGCTGGTCTATGAACTCAACCCCTTGGACCACCGAGGCCGGACCTT
GGAAATACCTGGCAACTCTGATCCCAATATGATCCCTGATGGGGACTTTAACAGCTACGTCAGG
GTTACAGCCTCAGATCCATTGGACACTTTAGGCTCTGAGGGGGCCTTGTACCAGGAGGCGT
20 GGCCTCCCTCTTGAGGCTTCCTCGAGGCTGTGGGGAGCAAACCATGATCTACTTGGCTCCGA
CACTGGCTGCTTCCCGCTACCTGGACAAGACAGAGCAGTGGAGCACACTGCCTCCCGAGAC
CAAGGACCACGCCGTGGATCTGATCCAGAAAGGCTACATGCGGATCCAGCAGTTTCGGAAGG
CGGATGGTTCCTATGCGGCTTGGTTGTCACGGGACAGCAGCACCTGGCTCACAGCCTTTGTG
TTGAAGGTCCTGAGTTTGGCCCAGGAGCAGGTAGGAGGCTCGCCTGAGAACTGCAGGAGA
25 CATCTAACTGGCTTCTGTCCCAGCAGCAGGCTGACGGCTCGTTCCAGGACCCCTGTCCAGTG
TTAGACAGGAGCATGCAGGGGGGTTTGGTGGGCAATGATGAGACTGTGGCACTCACAGCCTT
TGTGACCATCGCCCTTCATCATGGGCTGGCCGTCTTCCAGGATGAGGGTGCAGAGCCATTGA
AGCAGAGAGTGGAAGCCTCCATCTCAAAGGCAAACCTCATTTTTGGGGGAGAAAGCAAGTGCT
GGGCTCCTGGGTGCCACGCAGCTGCCATCACGGCCTATGCCCTGACACTGACCAAGGCGC
30 CTGTGGACCTGCTCGGTGTTGCCACAACAACCTCATGGCAATGGCCCAGGAGACTGGAGAT
AACCTGTACTGGGGCTCAGTCACTGGTTCTCAGAGCAATGCCGTGTGCCCCACCCCGGCTCC
TCGCAACCCATCCGACCCCATGCCCCAGGCCCCAGCCCTGTGGATTGAAACCACAGCCTACG
CCCTGCTGCACCTCCTGCTTCACGAGGGCAAAGCAGAGATGGCAGACCAGGCTTCGGCCTG
GCTCACCCGTCAGGGCAGCTTCCAAGGGGGATTCCGCAGTACCCAAGACACGGTGATTGCCC
35 TGGATGCCCTGTCTGCCTACTGGATTGCCTCCCACACCACTGAGGAGAGGGGTCTCAATGTG
ACTCTCAGCTCCACAGGCCGGAATGGGTTCAAGTCCCACGCGCTGCAGCTGAACAACCGCCA
GATTCGCGGCCTGGAGGAGGAGCTGCAGTTTTCTTGGGCAGCAAGATCAATGTGAAGGTGG

GAGGAAACAGCAAAGGAACCCTGAAGGTCCTTCGTACCTACAATGTCCTGGACATGAAGAACA
CGACCTGCCAGGACCTACAGATAGAAGTGACAGTCAAAGGCCACGTCGAGTACACGATGGAA
GCAAACGAGGACTATGAGGACTATGAGTACGATGAGCTTCCAGCCAAGGATGACCCAGATGCC
CCTCTGCAGCCCGTGACACCCCTGCAGCTGTTTGAGGGTCGGAGGAACCGCCGCAGGAGGG
5 AGGCGCCCAAGGTGGTGGAGGAGCAGGAGTCCAGGGTGCACTACACCGTGTGCATCTGGCG
GAACGGCAAGGTGGGGCTGTCTGGCATGGCCATCGCGGACGTCACCCTCCTGAGTGGATTC
CACGCCCTGCGTGCTGACCTGGAGAAGCTGACCTCCCTCTCTGACCGTTACGTGAGTCACTT
TGAGACCGAGGGGGCCCCACGTCCTGCTGTATTTTGA CT CGGTCCCCACCTCCCGGGAGTGCG
TGGGCTTTGAGGCTGTGCAGGAAGTGCCGGTGGGGCTGGTGCAGCCGGCCAGCGCAACCC
10 TGTACGACTACTACAACCCCGAGCGCAGATGTTCTGTGTTTTACGGGGCACCAAGTAAGAGCA
GACTCTTGCCACCTTGTGTTCTGCTGAAGTCTGCCAGTGTGCTGAGGGGAAGTGCCCTCGC
CAGCGTCGCGCCCTGGAGCGGGGTCTGCAGGACGAGGATGGCTACAGGATGAAGTTTGCCT
GCTACTACCCCGTGTGGAGTACGGCTTCCAGGTTAAGGTTCTCCGAGAAGACAGCAGAGCT
GCTTTCCGCCTCTTTGAGACCAAGATCACCCAAGTCCTGCACTTCACCAAGGATGTCAAGGCC
15 GCTGCTAATCAGATGCGCAACTTCCTGGTTCGAGCCTCCTGCCGCCTTCGCTTGAACCTGG
GAAAGAATATTTGATCATGGGTCTGGATGGGGCCACCTATGACCTCGAGGGACACCCCCAGTA
CCTGCTGGACTCGAATAGCTGGATCGAGGAGATGCCCTCTGAACGCCTGTGCCGGAGCACCC
GCCAGCGGGCAGCCTGTGCCCAGCTCAACGACTTCCTCCAGGAGTATGGCACTCAGGGGTG
CCAGGTGTGAGGGCTGCCCTCCACCTCCGCTGGGAGGAACCTGAACCTGGGAACCATGAA
20 GCTGGAAGCACTGCTGTGTCCGCTTTCATGAACACAGCCTGGGACCAGGGCATATTAAAGGCT
TTTGGCAGCAAAGTGTCAGTGTTGGC

SEQ ID NO:4:

AGAAGGTAGCAGACAGACAGACGGATCTAACCTCTCTTGGATCCTCCAGCCATGAGGCTGCTC
TGGGGGCTGATCTGGGCATCCAGCTTCTTCACCTTATCTCTGCAGAAGCCCAGGTTGCTCTTG
25 TTCTCTCCTTCTGTGGTTCATCTGGGGGTCCCCCTATCGGTGGGGGTGCAGCTCCAGGATGT
GCCCCGAGGACAGGTAGTGAAAGGATCAGTGTTCTGAGAAACCCATCTCGTAATAATGTCCC
CTGCTCCCCAAAGGTGGACTTCACCCTTAGCTCAGAAAGAGACTTCGCACTCCTCAGTCTCCA
GGTGCCCTTGAAAGATGCGAAGAGCTGTGGCCTCCATCAACTCCTCAGAGGCCCTGAGGTCC
AGCTGGTGGCCCATTCGCCATGGCTAAAGGACTCTCTGTCCAGAACGACAAACATCCAGGGTA
30 TCAACCTGCTCTTCTCCTCTCGCCGGGGGCACCTCTTTTTGCAGACGGACCAGCCCATTAC
ACCCTGGCCAGCGGGTTCGGTACCGGGTCTTTGCTCTGGATCAGAAGATGCGCCCCGAGCACT
GACACCATCACAGTCATGGTGGAGAACTCTCACGGCCTCCGCGTGCGGAAGAAGGAGGTGTA
CATGCCCTCGTCCATCTTCAGGATGACTTTGTGATCCCAGACATCTCAGAGCCAGGGACCTG
GAAGATCTCAGCCCGATTCTCAGATGGCCTGGAATCCAACAGCAGCACCCAGTTTGAGGTGAA
35 GAAATATGTCCTTCCCAACTTTGAGGTGAAGATCACCCCTGGAAAGCCCTACATCCTGACGGT
GCCAGGCCATCTTGATGAAATGCAGTTAGACATCCAGGCCAGGTACATCTATGGGAAGCCAGT

GCAGGGGGTGGCATATGTGCGCTTTGGGCTCCTAGATGAGGATGGTAAGAAGACTTTCTTTTCG
GGGGCTGGAGAGTCAGACCAAGCTGGTGAATGGACAGAGCCACATTTCCCTCTCAAAGGCAG
AGTTCCAGGACGCCCTGGAGAAGCTGAATATGGGCATTACTGACCTCCAGGGGCTGCGCCTC
TACGTTGCTGCAGCCATCATTGAGTCTCCAGGTGGGGAGATGGAGGAGGCAGAGCTCACATC
5 CTGGTATTTTGTGTCATCTCCCTTCTCCTTGGATCTTAGCAAGACCAAGCGACACCTTGTGCCT
GGGGCCCCCTTCTGCTGCAGGCCTTGGTCCGTGAGATGTCAGGCTCCCCAGCTTCTGGCAT
TCCTGTCAAAGTTTCTGCCACGGTGTCTTCTCCTGGGTCTGTTCTGAAGTCCAGGACATTCA
GCAAAACACAGACGGGAGCGGCCAAGTCAGCATTCCAATAATTATCCCTCAGACCATCTCAGA
GCTGCAGCTCTCAGTATCTGCAGGCTCCCCACATCCAGCGATAGCCAGGCTCACTGTGGCAG
10 CCCCACCTTCAGGAGGCCCCGGGTTTCTGTCTATTGAGCGGCCGGATTCTCGACCTCCTCGT
GTTGGGGACACTCTGAACCTGAACCTTGCAGAGCCGTGGGCAGTGGGGCCACCTTTTCTCATT
CTACTACATGATCCTATCCCGAGGGCAGATCGTGTTTCATGAATCGAGAGCCCAAGAGGACCT
GACCTCGGTCTCGGTGTTTGTGGACCATCACCTGGCACCCCTCCTTCTACTTTGTGGCCTTCTA
CTACCATGGAGACCACCCAGTGGCCAACTCCCTGCGAGTGGATGTCCAGGCTGGGGCCTGC
15 GAGGGCAAGCTGGAGCTCAGCGTGGACGGTGCCAAGCAGTACCGGAACGGGGAGTCCGTG
AAGCTCCACTTAGAAACCGACTCCCTAGCCCTGGTGGCGCTGGGAGCCTTGGACACAGCTCT
GTATGCTGCAGGCAGCAAGTCCCACAAGCCCCTCAACATGGGCAAGGTCTTTGAAGCTATGAA
CAGCTATGACCTCGGCTGTGGTCTGGGGGTGGGGACAGTGCCCTTCAGGTGTTCCAGGCA
GCGGGCCTGGCCTTTTCTGATGGAGACCAGTGGACCTTATCCAGAAAGAGACTAAGCTGTCC
20 CAAGGAGAAGACAACCCGAAAAAGAGAAACGTGAACTTCCAAAAGGCGATTAATGAGAAATT
GGGTCAGTATGCTTCCCCGACAGCCAAGCGCTGCTGCCAGGATGGGGTGACACGTCTGCCCA
TGATGCGTTCTGCGAGCAGCGGGCAGCCCGCGTGACAGAGCCGGACTGCCGGGAGCCCT
TCCTGTCCTGCTGCCAATTTGCTGAGAGTCTGCGCAAGAAGAGCAGGGACAAGGGCCAGGC
GGGCCTCCAACGAGCCCTGGAGATCCTGCAGGAGGAGGACCTGATTGATGAGGATGACATTC
25 CCGTGCGCAGCTTCTTCCCAGAGAACTGGCTCTGGAGAGTGGAAACAGTGGACCGCTTTCAA
ATATTGACACTGTGGCTCCCCGACTCTCTGACCACGTGGGAGATCCATGGCCTGAGCCTGTCC
AAAACCAAAGGCCTATGTGTGGCCACCCCAGTCCAGCTCCGGGTGTTCCGCGAGTTCCACCT
GCACCTCCGCCTGCCCATGTCTGTCCGCCGCTTTGAGCAGCTGGAGCTGCGGCCTGTCCTCT
ATAACTACCTGGATAAAAACCTGACTGTGAGCGTCCACGTGTCCCCAGTGGAGGGGGCTGTGC
30 CTGGCTGGGGGCGGAGGGCTGGCCCAGCAGGTGCTGGTGCCTGCGGGCTCTGCCCGGCCT
GTTGCCTTCTCTGTGGTGCCACGGCAGCCGCCGCTGTGTCTCTGAAGGTGGTGGCTCGAG
GGTCCTTCGAATTCCTGTGGGAGATGCGGTGTCCAAGGTTCTGCAGATTGAGAAGGAAGGG
GCCATCCATAGAGAGGAGCTGGTCTATGAACTCAACCCCTTGGACCACCGAGGCCGGACCTT
GGAAATACCTGGCAACTCTGATCCCAATATGATCCCTGATGGGGACTTTAACAGCTACGTCAGG
35 GTTACAGCCTCAGATCCATTGGACACTTTAGGCTCTGAGGGGGCCTTGTACCAGGAGGCGT
GGCCTCCCTCTTGAGGCTTCCTCGAGGCTGTGGGGAGCAAACCATGATCTACTTGGCTCCGA
CACTGGCTGCTTCCCGCTACCTGGACAAGACAGAGCAGTGGAGCACACTGCCTCCCGAGAC

CAAGGACCACGCCGTGGATCTGATCCAGAAAGGCTACATGCGGATCCAGCAGTTTTCGGAAGG
CGGATGGTTCCTATGCGGCTTGGTTGTCACGGGACAGCAGCACCTGGCTCACAGCCTTTGTG
TTGAAGGTCCTGAGTTTGGCCCAGGAGCAGGTAGGAGGCTCGCCTGAGAACTGCAGGAGA
CATCTAACTGGCTTCTGTCCCAGCAGCAGGCTGACGGCTCGTTCCAGGACCCCTGTCCAGTG
5 TTAGACAGGAGCATGCAGGGGGGTTTGGTGGGCAATGATGAGACTGTGGCACTCACAGCCTT
TGTGACCATCGCCCTTCATCATGGGCTGGCCGTCTTCCAGGATGAGGGTGCAGAGCCATTGA
AGCAGAGAGTGGAAGCCTCCATCTCAAAGGCCAACTCATTTTTGGGGGAGAAAGCAAGTGCT
GGGCTCCTGGGTGCCACGCAGCTGCCATCACGGCCTATGCCCTGACACTGACCAAGGCGC
CTGTGGACCTGCTCGGTGTTGCCACAACAACCTCATGGCAATGGCCCAGGAGACTGGAGAT
10 AACCTGTACTGGGGCTCAGTCACTGGTTCTCAGAGCAATGCCGTGTGCCCCACCCCGGCTCC
TCGCAACCCATCCGACCCCATGCCCCAGGCCCCAGCCCTGTGGATTGAAACCACAGCCTACG
CCCTGCTGCACCTCCTGCTTCACGAGGGGCAAAGCAGAGATGGCAGACCAGGCTTCGGCCTG
GCTCACCCGTCAGGGCAGCTTCCAAGGGGGATTCCGCAGTACCCAAGACACGGTGATTGCC
TGGATGCCCTGTCTGCCTACTGGATTGCCTCCCACACCACTGAGGAGAGGGGTCTCAATGTG
15 ACTCTCAGCTCCACAGGCCGGAATGGGTTCAAGTCCACGCGCTGCAGCTGAACAACCGCCA
GATTCGCGGCCTGGAGGAGGAGCTGCAGTTTTCTTGGGCAGCAAGATCAATGTGAAGGTGG
GAGGAAACAGCAAAGGAACCCTGAAGGTCCTTCGTACCTACAATGTCCTGGACATGAAGAACA
CGACCTGCCAGGACCTACAGATAGAAGTGACAGTCAAAGGCCACGTGAGTACACGATGGAA
GCAAACGAGGACTATGAGGACTATGAGTACGATGAGCTTCCAGCCAAGGATGACCCAGATGCC
20 CCTCTGCAGCCCGTGACACCCCTGCAGCTGTTTGAGGGTTCGGAGGAACCGCCGAGGAGGG
AGGCGCCCAAGCTGACCTCCCTCTCTGACCGTTACGTGAGTCACTTTGAGACCGAGGGGCCC
CACGTCTGTGTATTTTGA CTGGTCCCCACCTCCCGGGAGTGCGTGGGCTTTGAGGCTGT
GCAGGAAGTGCCGGTGGGGCTGGTGCAGCCGGCCAGCGCAACCCTGTACGACTACTACAAC
CCCGAGCGCAGATGTTCTGTGTTTTACGGGGCACCAAGTAAGAGCAGACTCTTGGCCACCTT
25 GTGTTCTGCTGAAGTCTGCCAGTGTGCTGAGGGGAAGTGCCCTCGCCAGCGTCGCGCCCTG
GAGCGGGGTCTGCAGGACGAGGATGGCTACAGGATGAAGTTTGCCTGCTACTACCCCGTGT
GGAGTACGGCTTCCAGGTTAAGGTTCTCCGAGAAGACAGCAGAGCTGCTTTCCGCCTCTTTG
AGACCAAGATCACCCAAGTCCTGCACTTCACCAAGGATGTCAAGGCCGCTGCTAATCAGATGC
GCAACTTCCTGGTTCGAGCCTCCTGCCGCTTCGCTTGGAACCTGGGAAAGAATATTTGATCA
30 TGGGTCTGGATGGGGCCACCTATGACCTCGAGGGACACCCCCAGTACCTGCTGGACTCGAAT
AGCTGGATCGAGGAGATGCCCTCTGAACGCCTGTGCCGGAGCACCCGCCAGCGGGCAGCCT
GTGCCCAGCTCAACGACTTCCTCCAGGAGTATGGCACTCAGGGGTGCCAGGTGTGAGGGCT
GCCCTCCACCTCCGCTGGGAGGAACCTGAACCTGGGAACCATGAAGCTGGAAGCACTGCT
GTGTCCGCTTTCATGAACACAGCCTGGGACCAGGGCATATTAAAGGCTTTTGGCAGCAAAGTG
35 TCAGTGTTGGC

SEQ ID NO:5:

MFSMRIVCLVLSVVGTAWTADSGEGDFLAEGGGVRGPRVVERHQSACKDSDWPFCSDWDWNYK
CPSGCRMKGLIDEVNQDFTNRINKLKNLSFEYQKNNKDSHSLTTNIMEILRGDFSSANNRDNTYNR
VSEDLRISRIEVLKRKVIEKVQHIQLLQKNVRAQLVDMKRLEVDIDIKIRSCRGSCSRALAREVDLKD
YEDQQKQLEQVIAKDLLPSRDRQHLLPLIKMKPVPDLVPGNFKSQQLQKVPPEWKALTDMPQMRMEL
5 ERPGGNEITRGGSTSYGTGSETESPRNPSSAGSWNSGSSGPGSTGNRNPSSGTGGTATWKP
GSSGPGSTGSWNSGSSGTGSTGNQNPSPRPGSTGTWNPSSSERGSAGHWTSESSVSGSTG
QWHSESGSFRPDSPGSGNARPNPDWGTFFEEVSGNVSPGTRREYHTEKLVTSKGDKEKELRTGKE
KVTSGSTTTTTRRSCSKTVTKTVIGPDGHKEVTKEVVTSEDGSDCPEAMDGLTSLGIGTLDGFRHR
HPDEAAFFDTASTGKTFPGFFSPMLGEFVSETESRGSESGIFTNTKESSSHHPGIAEFPSRGKSSS
10 YSKQFTSSTS YNRGDSTFESKSYKMADEAGSEADHEGTHSTKRGHAKSRPVRDCDDVLQTHPS
GTQSGIFNIKLPGSSKIFSVYCDQETSLGGWLLIQQRMDGSLNFNRTWQDYKRGFGSLNDEGE
FWLGNDYLHLLTQRGSVLRVELEDWAGNEAYAEYHFRVGSEAEGYALQVSSYEGTAGDALIEGSV
EEGAEYTSNNMNFSTFDRDADQWEENCAEVYGGGWWYNNCQAANLNGIYYPGGSYDPRNNS
PYEIENGVVVWSFRGADYSLRAVRMKIRPLVTQ

15 SEQ ID NO:6:

MFSMRIVCLVLSVVGTAWTADSGEGDFLAEGGGVRGPRVVERHQSACKDSDWPFCSDWDWNYK
CPSGCRMKGLIDEVNQDFTNRINKLKNLSFEYQKNNKDSHSLTTNIMEILRGDFSSANNRDNTYNR
VSEDLRISRIEVLKRKVIEKVQHIQLLQKNVRAQLVDMKRLEVDIDIKIRSCRGSCSRALAREVDLKD
YEDQQKQLEQVIAKDLLPSRDRQHLLPLIKMKPVPDLVPGNFKSQQLQKVPPEWKALTDMPQMRMEL
20 ERPGGNEITRGGSTSYGTGSETESPRNPSSAGSWNSGSSGPGSTGNRNPSSGTGGTATWKP
GSSGPGSTGSWNSGSSGTGSTGNQNPSPRPGSTGTWNPSSSERGSAGHWTSESSVSGSTG
QWHSESGSFRPDSPGSGNARPNPDWGTFFEEVSGNVSPGTRREYHTEKLVTSKGDKEKELRTGKE
KVTSGSTTTTTRRSCSKTVTKTVIGPDGHKEVTKEVVTSEDGSDCPEAMDGLTSLGIGTLDGFRHR
HPDEAAFFDTASTGKTFPGFFSPMLGEFVSETESRGSESGIFTNTKESSSHHPGIAEFPSRGKSSS
25 YSKQFTSSTS YNRGDSTFESKSYKMADEAGSEADHEGTHSTKRGHAKSRPVRGIHTSPLGKPSL
SP

SEQ ID NO:7:

AGCAATCCTTTCTTTCTAGCTGGAGTGCTCCTCAGGAGCCAGCCCCACCCTTAGAAAAGATGTT
TTCCATGAGGATCGTCTGCCTGGTCCTAAGTGTGGTGGGCACAGCATGGACTGCAGATAGTG
30 GTGAAGGTGACTTTCTAGCTGAAGGAGGAGGCGTGCGTGGCCCAAGGGTTGTGGAAAGACAT
CAATCTGCCTGCAAAGATTCAGACTGGCCCTTCTGCTCTGATGAAGACTGGAAC TACAAATGC
CCTTCTGGCTGCAGGATGAAAGGGTTGATTGATGAAGTCAATCAAGATTTTACAAACAGAATAA
ATAAGCTCAAAAATTCAC TATTTGAATATCAGAAGAACAATAAGGATTCTCATTGTTGACCACTA
ATATAATGGAAATTTTGAGAGGCGATTTTCTCAGCCAATAACCGTGATAATACCTACAACCGA
35 GTGTCAGAGGATCTGAGAAGCAGAATTGAAGTCCTGAAGCGCAAAGTCATAGAAAAAGTACAG

CATATCCAGCTTCTGCAGAAAAATGTTAGAGCTCAGTTGGTTGATATGAAACGACTGGAGGTGG
ACATTGATATTAAGATCCGATCTTGTGCGAGGGTCATGCAGTAGGGCTTTAGCTCGTGAAGTAGA
TCTGAAGGACTATGAAGATCAGCAGAAGCAACTTGAACAGGTCATTGCCAAAGACTTACTTCCC
TCTAGAGATAGGCAACACTTACCACTGATAAAAATGAAACCAGTTCCAGACTTGGTTCCCGGAA
5 ATTTTAAGAGCCAGCTTCAGAAGGTACCCCCAGAGTGGAAGGCATTAACAGACATGCCGCAGA
TGAGAATGGAGTTAGAGAGACCTGGTGGAAATGAGATTACTCGAGGAGGCTCCACCTCTTATG
GAACCGGATCAGAGACGGAAAGCCCCAGGAACCCTAGCAGTGCTGGAAGCTGGAACCTCTGG
GAGCTCTGGACCTGGAAGTACTGGAAACCGAAACCCTGGGAGCTCTGGGACTGGAGGGACT
GCAACCTGGAAACCTGGGAGCTCTGGACCTGGAAGTACTGGAAGCTGGAACCTCTGGGAGCT
10 CTGGAACCTGGAAGTACTGGAAACCAAACCCCTGGGAGCCCTAGACCTGGTAGTACCGGAACC
TGGAATCCTGGCAGCTCTGAACGCGGAAGTGCTGGGCACTGGACCTCTGAGAGCTCTGTATC
TGGTAGTACTGGACAATGGCACTCTGAATCTGGAAGTTTTAGGCCAGATAGCCCAGGCTCTGG
GAACGCGAGGCCTAACAACCCAGACTGGGGCACATTTGAAGAGGTGTCAGGAAATGTAAGTC
CAGGGACAAGGAGAGAGTACCACACAGAAAACTGGTCACTTCTAAAGGAGATAAAGAGCTCA
15 GGACTGGTAAAGAGAAGGTCACCTCTGGTAGCACAACCACCACGCGTCGTTTCATGCTCTAAAA
CCGTTACTAAGACTGTTATTGGTCCTGATGGTCACAAAGAAGTTACCAAAGAAGTGGTGACCTC
CGAAGATGGTTCTGACTGTCCCGAGGCAATGGATTTAGGCACATTGTCTGGCATAGGTACTCT
GGATGGGTTCCGCCATAGGCACCCTGATGAAGCTGCCTTCTTCGACACTGCCTCAACTGGAAA
AACATTCCCAGGTTTCTTCTCACCTATGTTAGGAGAGTTTGTCAGTGAGACTGAGTCTAGGGGC
20 TCAGAATCTGGCATCTTCACAAATACAAAGGAATCCAGTTCTCATCACCCCTGGGATAGCTGAAT
TCCCTTCCCGTGGTAAATCTTCAAGTTACAGCAAACAATTTACTAGTAGCACGAGTTACAACAGA
GGAGACTCCACATTTGAAAGCAAGAGCTATAAAATGGCAGATGAGGCCGGAAGTGAAGCCGAT
CATGAAGGAACACATAGCACCAAGAGAGGCCATGCTAAATCTCGCCCTGTCAGAGACTGTGAT
GATGTCTCCAAACACATCCTTCAGGTACCCAAAGTGGCATTTCATATCAAGCTACCGGGAT
25 CCAGTAAGATTTTTCTGTTTATTGCGATCAAGAGACCAGTTTGGGAGGATGGCTTTTGATCCA
GCAAAGAATGGATGGATCACTGAATTTTAACCGGACCTGGCAAGACTACAAGAGAGGTTTCGG
CAGCCTGAATGACGAGGGGGAAGGAGAATTCTGGCTAGGCAATGACTACCTCCACTTACTAAC
CCAAAGGGGCTCTGTTCTTAGGGTTGAATTAGAGGACTGGGCTGGGAATGAAGCTTATGCAGA
ATATCACTTCCGGGTAGGCTCTGAGGCTGAAGGCTATGCCCTCCAAGTCTCCTCCTATGAAGG
30 CACTGCGGGTGATGCTCTGATTGAGGGTTCCGTAGAGGAAGGGGCAGAGTACACCTCTCACA
ACAACATGCAGTTCAGCACCTTTGACAGGGATGCAGACCAGTGGGAAGAGAACTGTGCAGAA
GTCTATGGGGGAGGCTGGTGGTATAATAACTGCCAAGCAGCCAATCTCAATGGAATCTACTACC
CTGGGGGCTCCTATGACCCAAGGAATAACAGTCCTTATGAGATTGAGAATGGAGTGGTCTGGG
TTTCCTTTAGAGGGGCAGATTATTCCTCAGGGCTGTTTCGCATGAAAATTAGGCCCTTGTGAC
35 CCAATAGGCTGAAGAAGTGGGAATGGGAGCACTCTGTCTTCTTTGCTAGAGAAGTGGAGAGAA
AATACAAAAGGTAAAGCAGTTGAGATTCTCTACAACCTAAAAAATTCCTAGGTGCTATTTTCTTAT
CCTTTGTAAGTGTAGCTAAATGTACCTGAGACATATTAGTCTTTGAAAATAAAGTTATGTAAGTT

TTTTTTATCTTTAAATAGCTCTGTGGGTTTTAACATTTTTATAAAGATATACCAAGGGCCATTTCAGT
ACATCAGGAAAGTGGCAGACAGAAGCTTCTCTCTGCAACCTTGAAGACTATTGGTTTGAGAAC
TTCTCTTCCCATACCAACCCAAAATCATAATGCCATTGGAAAGCAAAAAGTTGTTTTATCCATTGGA
TTTGAATTGTTTTAAGCCAATATTTTAAGGTAAAACTCACTGAATCTAACCATAGCTGACCTTTGT
5 AGTAGAATTTACAACCTTATAATTACAATGCACAATTTATAATTACAATATGTATTTATGTCTTTTGCTA
TGGAGCAAATCCAGGAAGGCAAGAGAAACATTCTTTCCTAAATATAAATGAAAATCTATCCTTTA
AACTCTTCCACTAGACGTTGTAATGCACACTTATTTTTTTTCCCAAGGAGTAACCAATTTCTTTCTA
AAACACATTTAAAATTTTAAAACCTATTTATGAATATTAAAAAAAGACATAATTCACACATTAATAAAC
AATCTCCCAAGTATTGATTTAACTTCATTTTTCTAATAATCATAACTATATTCTGTGACATGCTAAT
10 TATTATTAAATGTAAGTCGTTAGTTCGAAAGCCTCTCACTAAGTATGATCTATGCTATATTCAAAAT
TCAACCCATTTACTTTGGTCAATATTTGATCTAAGTTGCATCTTTAATCCTGGTGGTCTTGCCTTC
TGATTTTTAATTTGTATCCTTTTTCTATTAAGATATATTTGTCATTTTCTCTTGAATATGTATTAAATA
TCCCAAGC

SEQ ID NO:8:

15 AGCAATCCTTTCTTTTCTAGCTGGAGTGCTCCTCAGGAGCCAGCCCCACCCTTAGAAAAGATGTT
TTCCATGAGGATCGTCTGCCTGGTCCTAAGTGTGGTGGGCACAGCATGGACTGCAGATAGTG
GTGAAGGTGACTTTCTAGCTGAAGGAGGAGGCGTGCGTGGCCCAAGGGTTGTGGAAAGACAT
CAATCTGCCTGCAAAGATTCAGACTGGCCCTTCTGCTCTGATGAAGACTGGAACCTACAAATGC
CCTTCTGGCTGCAGGATGAAAGGGTTGATTGATGAAGTCAATCAAGATTTTACAAACAGAATAA
20 ATAAGCTCAAAAATTCACCTATTTGAATATCAGAAGAACAATAAGGATTCTCATTTCGTTGACCACTA
ATATAATGGAAATTTTGAGAGGCGATTTTTCTCAGCCAATAACCGTGATAATACCTACAACCGA
GTGTCAGAGGATCTGAGAAGCAGAATTGAAGTCCTGAAGCGCAAAGTCATAGAAAAGTACAG
CATATCCAGCTTCTGCAGAAAAATGTTAGAGCTCAGTTGGTTGATATGAAACGACTGGAGGTGG
ACATTGATATTAAGATCCGATCTTGTGAGGGTCATGCAGTAGGGCTTTAGCTCGTGAAGTAGA
25 TCTGAAGGACTATGAAGATCAGCAGAAGCAACTTGAACAGGTCATTGCCAAAGACTTACTTCCC
TCTAGAGATAGGCAACACTTACCACTGATAAAAATGAAACCAGTTCCAGACTTGGTTCCCGGAA
ATTTTAAGAGCCAGCTTCAGAAGGTACCCCCAGAGTGGAAGGCATTAACAGACATGCCGCAGA
TGAGAATGGAGTTAGAGAGACCTGGTGGAAATGAGATTACTCGAGGAGGCTCCACCTCTTATG
GAACCGGATCAGAGACGGAAAGCCCCAGGAACCCTAGCAGTGCTGGAAGCTGGAACCTCTGG
30 GAGCTCTGGACCTGGAAGTACTGGAACCGAAACCCTGGGAGCTCTGGGACTGGAGGGACT
GCAACCTGGAACCTGGGAGCTCTGGACCTGGAAGTACTGGAAGCTGGAACCTCTGGGAGCT
CTGGAACCTGGAAGTACTGGAACCCAAAACCCTGGGAGCCCTAGACCTGGTAGTACCGGAACC
TGGAATCCTGGCAGCTCTGAACGCGGAAGTGCTGGGCACTGGACCTCTGAGAGCTCTGTATC
TGGTAGTACTGGACAATGGCACTCTGAATCTGGAAGTTTTAGGCCAGATAGCCCAGGCTCTGG
35 GAACGCGAGGCCTAACAACCCAGACTGGGGCACATTTGAAGAGGTGTCAGGAAATGTAAGTC
CAGGGACAAGGAGAGAGTACCACACAGAAAACTGGTCACTTCTAAAGGAGATAAAGAGCTCA

GGACTGGTAAAGAGAAGGTCACCTCTGGTAGCACAACCACCACGCGTCGTTTCATGCTCTAAAA
CCGTTACTAAGACTGTTATTGGTCCTGATGGTCACAAAGAAGTTACCAAAGAAGTGGTGACCTC
CGAAGATGGTTCTGACTGTCCCAGAGCAATGGATTTAGGCACATTGTCTGGCATAGGTACTCT
GGATGGGTTCCGCCATAGGCACCCTGATGAAGCTGCCTTCTTCGACACTGCCTCAACTGGAAA
5 AACATTCCCAGGTTTCTTCTCACCTATGTTAGGAGAGTTTGTCACTGAGACTGAGTCTAGGGGC
TCAGAATCTGGCATCTTCACAAATACAAAGGAATCCAGTTCTCATCACCTGGGATAGCTGAAT
TCCCTTCCCGTGGTAAATCTTCAAGTTACAGCAAACAATTTACTAGTAGCACGAGTTACAACAGA
GGAGACTCCACATTTGAAAGCAAGAGCTATAAAATGGCAGATGAGGCCGGAAGTGAAGCCGAT
CATGAAGGAACACATAGCACCAAGAGAGGCCATGCTAAATCTCGCCCTGTCAGAGGTATCCAC
10 ACTTCTCCTTTGGGGAAGCCTTCCCTGTCCCCCTAGACTAAGTTAAATATTTCTGCACAGTGTT
CCCATGGCCCCCTTGCAATTCCTTCTTAAGTCTCTGTTACACGTCATTGAAACTACACTTTTTTGG
TCTGTTTTTGTGCTAGACTGTAAGTTCCTTGGGGGCAGGGCCTTTGTCTGTCTCATCTCTGTAT
TCCCAAATGCCTAACAGTACAGAGCCATGACTCAATAAATACATGTAAATGGATGAATG

SEQ ID NO:9:

15 MKILILGIFLFLCSTPAWAKEKHYYIGIIEETTWDYASDHGEKKLISVDTEHSNIYLQNGPDRIGRLYKK
ALYLQYTDETFRTTIEKPVWLGLGPIIKAETGDKVYVHLKNLASRPYTFHSHGITYYKEHEGAIIYPD
NTTDFQRADDKVYPGEQYTYMLLATEEQSPGEGDGNCVTRIYHSHIDAPKDIASGLIGPLIICKKDS
LDKEKEKHIDREFVVMFSVVDENFSWYLEDNIKTYCSEPEKVDKDNEDFQESNRMYSVNGYTFGS
LPGLSMCAEDRVKWYLFMGNEVDVHAFFHGGQALTNKNYRIDTINLFPATLFDAYMVAQNPGEW
20 MLSCQNLNHLKAGLQAFFQVQECNKSSSKDNIRGKHVRHYIAAEEIIWNYAPSGIDIFTKENLTAP
GSDSAVFFEQGTTRIGGSYKKLVYREYTDASFTNRKERGPREEHLGILGPVIWAEVGDITRVTFHN
KGAYPLSIEPIGVRFNKNNEGTYYSNPYNPQSRSVPPSASHVAPTETFTYEWTVPKVGPNTADP
VCLAKMYYSAVDPTKDIFTGLIGPMKICKKGS L HANGRQKDVDKEFYLFPTVFDENESLLLEDNIRM
FTTAPDQVDKEDEDFQESNKMHS MNGFM YGNQ PGLTMCKGDSVWYLF SAGNEADVHGIYFSG
25 NTYLWRGERRDTANLFPQTS L THMWPDT EGT FNVECLT TDHYTGGMKQKYTVNQCRQSEDS
TFYLGERTYYIAAVEVEWDYSPQREWEKELHHLQEQNVSN AFLDKGEFYIGSKYKKVYRQYTD S
TFRVPVERKAEEEEHLGILGPQLHADVGDKVKIIFKNMATRPYSIHAHGVQTESSTVTPTLPGETLTY
VWKIPERSGAGTEDSACIPWAYYSTVDQVKDLYSGLIGPLIVCRRPYLKVFNPRRKLEFALLFLVFD
ENESWYLLDNKTYSDHPEKVNKDDEEFIESNKMHAINGRMFGNLQGLTMHVGDEVNWYLMGM
30 GNEIDLHTVHFHGHFSFYKHGRGVYSSDVFDIFPGTYQTLEMFPRTPGIWLLHCHVTDHIHAGMETT
YTVLQNE DTKSG

SEQ ID NO:10:

ACACCCTAATGCCTCCAACAATAACTGTTGACTTTTTATTTTCAGTCAGAGAAGCCTGGCAACC
AAGAACTGTTTTTTTGGTGGTTTACGAGAACTTAACTGAATTGGAAAATATTTGCTTTAATGAAAC
35 AATTTACTCTTGTGCAACACTAAATTGTGTCAATCAAGCAAATAAGGAAGAAAGTCTTATTTATAA

AATTGCCTGCTCCTGATTTTACTTCATTTCTTCTCAGGCTCCAAGAAGGGGAAAAAATGAAGA
TTTTGATACTTGGTATTTTTCTGTTTTATGTAGTACCCCAGCCTGGGCGAAAGAAAAGCATTATT
ACATTGGAATTATTGAAACGACTTGGGATTATGCCTCTGACCATGGGGAAAAAGAACTTATTTCT
GTTGACACGGAACATTCCAATATCTATCTTCAAATGGCCCAGATAGAATTGGGAGACTATATAA
5 GAAGGCCCTTTATCTTCAGTACACAGATGAAACCTTTAGGACAACCTATAGAAAAACCGGTCTGG
CTTGGGTTTTTATAGGCCCTATTATCAAAGCTGAAACTGGAGATAAAGTTTATGTACACTTAAAAAA
CCTTGCCTCTAGGCCCTACACCTTTCATTACATGGAATAACTTACTATAAGGAACATGAGGGG
GCCATCTACCCTGATAACACCACAGATTTTCAAAGAGCAGATGACAAAGTATATCCAGGAGAGC
AGTATACATACATGTTGCTTGCCACTGAAGAACAAAGTCCTGGGGAAGGAGATGGCAATTGTGT
10 GACTAGGATTTACCATTCCCACATTGATGCTCCAAAAGATATTGCCTCAGGACTCATCGGACCT
TTAATAATCTGTAAAAAGATTCTCTAGATAAAGAAAAAGAAAAACATATTGACCGAGAATTTGTG
GTGATGTTTTCTGTGGTGGATGAAAATTTTACAGCTGGTACCTAGAAGACAACATTAAACCTACT
GCTCAGAACCAGAGAAAGTTGACAAAGACAACGAAGACTTCCAGGAGAGTAACAGAATGTATT
CTGTGAATGGATACACTTTTGAAGTCTCCCAGGACTCTCCATGTGTGCTGAAGACAGAGTAA
15 AATGGTACCTTTTTGGTATGGGTAATGAAGTTGATGTGCACGCAGCTTTCTTTCACGGGCAAGC
ACTGACTAACAAGAACTACCGTATTGACACAATCAACCTCTTTCCTGCTACCCTGTTTGATGCTT
ATATGGTGGCCCAGAACCCTGGAGAATGGATGCTCAGCTGTCAGAATCTAAACCATCTGAAAG
CCGGTTTTGCAAGCCTTTTTCCAGGTCCAGGAGTGTAACAAGTCTTCATCAAAGGATAATATCCG
TGGGAAGCATGTTAGACACTACTACATTGCCGCTGAGGAAATCATCTGGAACATGCTCCCTCT
20 GGTATAGACATCTTCACTAAAGAAAACCTTAACAGCACCTGGAAGTGAAGTCAAGCGGTGTTTTTTG
AACAAGGTACCACAAGAATTGGAGGCTCTTATAAAAAGCTGGTTTATCGTGAGTACACAGATGC
CTCCTTCACAAATCGAAAGGAGAGAGGCCCTGAAGAAGAGCATCTTGGCATCCTGGGTCTCTG
TCATTTGGGCAGAGGTGGGAGACACCATCAGAGTAACCTTCCATAACAAAGGAGCATATCCCC
TCAGTATTGAGCCGATTGGGGTGAGATTCAATAAGAACAACGAGGGGCACATACTATTCCCCAAA
25 TTACAACCCCCAGAGCAGAAGTGTGCCTCCTTCAGCCTCCCATGTGGCACCCACAGAAACATT
CACCTATGAATGGACTGTCCCCAAAGAAGTAGGACCCACTAATGCAGATCCTGTGTGTCTAGCT
AAGATGTATTATTCTGCTGTGGATCCCCTAAAGATATATTCACTGGGCTTATTGGGCCAATGAA
AATATGCAAGAAAGGAAGTTTACATGCAAATGGGAGACAGAAAGATGTAGACAAGGAATTCTAT
TTGTTTCCTACAGTATTTGATGAGAATGAGAGTTTACTCCTGGAAGATAATATTAGAATGTTTACA
30 ACTGCACCTGATCAGGTGGATAAGGAAGATGAAGACTTTCAGGAATCTAATAAAATGCACTCCA
TGAATGGATTCATGTATGGGAATCAGCCGGGTCTCACTATGTGCAAAGGAGATTCCGGTCGTGT
GGTACTTATTCAGCGCCGGAATGAGGCCGATGTACATGGAATATACTTTTCAGGAAACACATAT
CTGTGGAGAGGAGAACGGAGAGACACAGCAAACCTCTTCCCTCAAACAAGTCTTACGCTCCA
CATGTGGCCTGACACAGAGGGGACTTTTAATGTTGAATGCCTTACAACCTGATCATTACACAGGC
35 GGCATGAAGCAAAAATATACTGTGAACCAATGCAGGCGGCAGTCTGAGGATTCCACCTTCTAC
CTGGGAGAGAGGACATACTATATCGCAGCAGTGGAGGTGGAATGGGATTATTTCCCAACAAAGG
GAGTGGGAAAAGGAGCTGCATCATTTACAAGAGCAGAATGTTTCAAATGCATTTTTAGATAAGG

GAGAGTTTTACATAGGCTCAAAGTACAAGAAAGTTGTGTATCGGCAGTATACTGATAGCACATTC
CGTGTTCCAGTGGAGAGAAAAGCTGAAGAAGAACATCTGGGAATTCTAGGTCCACAACCTTCAT
GCAGATGTTGGAGACAAAGTCAAATTATCTTTAAAAACATGGCCACAAGGCCCTACTCAATAC
ATGCCCATGGGGTACAAACAGAGAGTTCTACAGTTACTCCAACATTACCAGGTGAAACTCTCAC
5 TTACGTATGGAAAATCCCAGAAAGATCTGGAGCTGGAACAGAGGATTCTGCTTGTATTCCATGG
GCTTATTATTCAACTGTGGATCAAGTTAAGGACCTCTACAGTGGATTAATTGGCCCCCTGATTGT
TTGTCTGAAGACCTTACTTGAAAGTATTCAATCCCAGAAGGAAACTGGAATTTGCCCTTCTGTTT
CTAGTTTTTTGATGAGAATGAATCTTGGTACTTAGATGACAACATCAAAACATACTCTGATCACCC
CGAGAAAAGTAAACAAAGATGATGAGGAATTCATAGAAAGCAATAAAATGCATGCTATTAATGGAA
10 GAATGTTTGGAAACCTACAAGGCCTCACAATGCACGTGGGAGATGAAGTCAACTGGTATCTGA
TGGGAATGGGCAATGAAATAGACTTACACACTGTACATTTTTCACGGCCATAGCTTCCAATACAA
GCACAGGGGAGTTTATAGTTCTGATGTCTTTGACATTTTCCCTGGAACATACCAAACCCTAGAA
ATGTTTCCAAGAACACCTGGAATTTGGTACTCCACTGCCATGTGACCGACCACATTCATGCTG
GAATGGAAACCACTTACACCGTTCTACAAAATGAAGACACCAAATCTGGCTGAATGAAATAAATT
15 GGTGATAAGTGGAAAAAAGAGAAAAACCAATGATTCATAACAATGTATGTGAAAGTGTAATAAG
AATGTTACTTTGGAATGACTATAAACATTAAAAGAAGACTGGAAGCATACAACCTTTGTACATTTGT
GGGGGAAAACCTATTAATTTTTTTGCAAATGGAAAGATCAACAGACTATATAATGATACATGACTGAC
ACTTGTACACTAGGTAATAAAACTGATTCATACAGTCTAATGATATCACCGCTGTTAGGGTTTTAT
AAAACCTGCATTTAAAAAAGATCTATGACCAGATATTCTCCTGGGTGCTCCTCAAAGGAACACTA
20 TTAAGGTTTCATTGAAATGTTTTCAATCATTGCCTTCCCATTGATCCTTCTAACATGCTGTTGACAT
CACACCTAATATTCAGAGGGAATGGGCAAGGTATGAGGGAAGGAAATAAAAAATAAAATAATAA
AATAGAATGACACAAATTTGAGTTTTGTGAACCCCTGAACAGATGGTCTTAAGGACGTTATCTG
GAACTGGAGAAAAGCAGAGTTGAGAGACAATTCTATAGATTAAATCCTGGTAAGGACAAACATT
GCCATTAGAAGAAAAGCTTCAAATAGACCTGTGGCAGATGTCACATGAGTAGAATTTCTGCCC
25 AGCCTTAACCTGCATTCAGAGGATAATATCAATGAACTAACTTGAACATAAAATTTTTTAAACAAA
AAGTTATAAATGAAGACACATGGTTGTGAATACAATGATGTATTTCTTTATTTTACATACACTCTA
GCTAAAAGAGCAAGAGTACACATCAACAAAAATGGAAACAAGGCTTTGGCTGAAAAAACATG
CATTTGACAAATCATGTTAATAGCTAGACAAGAAGAAAGTTAGCTTTGTAACTTCTACTTCATTT
GATTCAGAGAAACAGAGCATGAGTTTTCTTAAAAGTAACAAGAAAAGGAACAAAAAAATGAGG
30 TTTGAAATCTTTTACCATGGCAAAACATTAACATCTTTCTCAAAAACATAGAGAAATCTGGAAAAA
TCAAGAAGATAAAATTCTGGACCAGTTAGTGACATTCTTTCAAGCATACTTGTAATGTTTCCTT
AAAGTGTTCTTGGGATGAAAATGATTGTCTGTCTCCAACAACAGTGAACCTGATGTTGTTCTT
GGAATAAAAGTCAATCCCCACCTTAAAAAAAAAAAAAA

SEQ ID NO:11:

35 MAKLIALTLLGMGLALFRNHQSSYQTRLNALREVQPVLPNCNLVKGIETGSEDLEILPNGLAFISS
GLKYPGIKSFNPNSPGKILLMDLNEEDPTVLELGITGSKFDVSSFNPHGISTFTDEDNAMYLLVNVH

PDAKSTVELFKFQEEEEKSLLHLKTIRHKLLPNLNDIVAVGPEHFYGTNDHYFLDPYLQSWEMYLGLA
WSYVVYYSPSEVRVVAEGFDFANGINISPDGKYVYIAELLAHKIHVYEKHANDTLPLKSLDFNTLV
DNISVDPETGDLWVGCHPNGMKIFFYDSENPPASEVLRIQNILTEEPKVTQVYAENGTVLQGSTVA
SVYKGKLLIGTVFHKALYCEL

5 SEQ ID NO:12:

AATCGGCGCTGCCCCAGCAGGGCTGCGGCTGCAGGCAGGCAGAGCCTCCTAGCCCGTCGGT
GTCTGCGCCCATCGATCCCTTTGTCTATCCCCGACCATGGCGAAGCTGATTGCGCTCACCCCTC
TTGGGGATGGGACTGGCACTCTTCAGGAACCACCAGTCTTCTTACCAAACACGACTTAATGCT
CTCCGAGAGGTACAACCCGTAGAACTTCCTAACTGTAATTTAGTTAAAGGAATCGAACTGGCT
10 CTGAAGACTTGGAGATACTGCCTAATGGACTGGCTTTTCATTAGCTCTGGATTAAAGTATCCTGG
AATAAAGAGCTTCAACCCCAACAGTCCTGGAAAAATACTTCTGATGGACCTGAATGAAGAAGAT
CCAACAGTGTTGGAATTGGGGATCACTGGAAGTAAATTTGATGTATCTTCATTTAACCCTCATGG
GATTAGCACATTCACAGATGAAGATAATGCCATGTACCTCCTGGTGGTGAACCATCCAGATGCC
AAGTCCACAGTGGAGTTGTTTAAATTTCAAGAAGAAGAAAAATCGCTTTTGCATCTAAAAACCAT
15 CAGACATAAACTTCTGCCTAATTTGAATGATATTGTTGCTGTGGGACCTGAGCACTTTTATGGCA
CAAATGATCACTATTTTCTTGACCCCTACTTACAATCCTGGGAGATGTATTTGGGTTTAGCGTGG
TCGTATGTTGTCTACTATAGTCCAAGTGAAGTTCGAGTGGTGGCAGAAGGATTTGATTTTGCTAA
TGGAATCAACATTTACCCGATGGCAAGTATGTCTATATAGCTGAGTTGCTGGCTCATAAGATTC
ATGTGTATGAAAAGCATGCTAATTGGACTTTAACTCCATTGAAGTCCCTTGACTTTAATACCCTC
20 GTGGATAACATATCTGTGGATCCTGAGACAGGAGACCTTTGGGTTGGATGCCATCCCAATGGC
ATGAAAATCTTCTTCTATGACTCAGAGAATCCTCCTGCATCAGAGGTGCTTCGAATCCAGAACA
TTCTAACAGAAGAACCTAAAGTGACACAGGTTTATGCAGAAAATGGCACAGTGTTGCAAGGCA
GTACAGTTGCCTCTGTGTACAAAGGGAAACTGCTGATTGGCACAGTGTTTCACAAAGCTCTTTA
CTGTGAGCTCTAACAGACCGATTTGCACCCATGCCATAGAACTGAGGCCATTATTTCAACCGC
25 TTGCCATATTCCGAGGACCCAGTGTTCTTAGCTGAACAATGAATGCTGACCCTAAATGTGGACA
TCATGAAGCATCAAAGCACTGTTTAACTGGGAGTGATATGATGTGTAGGGCTTTTTTTTGAGAAT
ACACTATCAAATCAGTCTTGGAATACTTGAAAACCTCATTTACCATAAAAATCCTTCTCACTAAAA
TGGATAAATCAGTTATGTCAATTGTCAGATATTAATAACAGTGTGTGACCCCAAAAGTACTTACC
CTAAACATGTGTTGCCTGGAAGCACATGTGTGTATCGCTGCCTTGCCATGTCTTGTTGAGAAG
30 ACACAGGGGAGCAGGGTTAGCTCACGTGTCTTTAGAACTCCAGTACTCACCCAGGGACTCCA
GTTTCACAGGCCAGAAAACATATGCATTATGAAGTTCCCCTCTACTCCATGCACATAGTAAGTCTG
ACTATGGCAGTCAGACTTACTTACTCCCATTTTCCCTTCGATATATGACTTTTTCTCAGTAAATATT
AACCTGAATTATTCCAAAAA

En otra realización preferida de este aspecto de la invención, la detección de la cantidad de
35 expresión de cualquiera de los genes C4A, FGA, CP y PON1 se realiza mediante:

i. un procedimiento de perfil genético, tal como una micromatriz y/o un procedimiento que comprende PCR, tal como una PCR en tiempo real, transferencia Northern y/o

iii. un procedimiento inmunológico y/o inmunoensayo, ELISA y/o transferencia Western

5 En otra realización preferida, la detección de los niveles de expresión de los genes se realiza mediante Q-RT-PCR.

En otra realización preferida de este aspecto de la invención, la detección de la cantidad de expresión de cualquiera de los genes C4A, FGA, CP y PON1 se realiza mediante un inmunoensayo.

10 El término “inmunoensayo”, tal y como se utiliza en la presente descripción se refiere a cualquier técnica analítica que se basa en la reacción de la conjugación de un anticuerpo con un antígeno. Ejemplos de inmunoensayos conocidos en el estado de la técnica son, por ejemplo, pero sin limitarse: inmunoblot, ensayo inmunoabsorbente ligado a enzimas (ELISA), inmunoensayo lineal (LIA), radioinmunoensayo (RIA), inmunofluorescencia, x-map o chips de proteína.

15 En otra realización preferida, el inmunoensayo es un ensayo inmunoabsorbente ligado a enzimas o ELISA (Enzyme-Linked ImmunoSorbent Assay). El ELISA se basa en la premisa de que un inmunorreactivo (antígeno o anticuerpo) puede ser inmovilizado en un soporte sólido, poniendo luego ese sistema en contacto con una fase fluida que contiene el reactivo complementario que puede unirse a un compuesto marcador. Existen diferentes tipos de ELISA: ELISA directo, ELISA indirecto o ELISA sándwich.

20 El término “compuesto marcador”, tal y como se utiliza en la presente descripción, se refiere a un compuesto capaz de dar lugar a una señal cromogénica, fluorogénica, radiactiva y/o quimioluminiscente que permita la detección y cuantificación de la cantidad de anticuerpos frente a C4A, FGA, CP y PON1. El compuesto marcador se selecciona de la lista que comprende radioisótopos, enzimas, fluoróforos o cualquier molécula susceptible de ser conjugada con otra
25 molécula o detectada y/o cuantificada de forma directa. Este compuesto marcador puede unirse al anticuerpo directamente, o a través de otro compuesto. Algunos ejemplos de compuestos marcadores que se unen directamente son, pero sin limitarse, enzimas como la fosfatasa alcalina o la peroxidasa, isótopos radiactivos como ^{32}P o ^{35}S , fluorocromos como fluoresceína o partículas metálicas, para su detección directa mediante colorimetría, auto-radiografía,
30 fluorimetría, o metalografía respectivamente.

Un **aspecto G** de la invención se refiere a un kit o dispositivo que comprende los elementos necesarios para cuantificar la cantidad de expresión de los genes que se seleccionan de entre *C4A*, *CP* y *PON1*, o cualquiera de sus combinaciones.

5 Aún más preferiblemente, el kit de este aspecto de la invención comprende los anticuerpos que se seleccionan de la lista que consiste en anticuerpos: anti-*C4A*, anti-*CP* y anti- *PON1*, o cualquiera de sus combinaciones. Preferiblemente, el kit de la invención comprende simultáneamente al menos un anticuerpo de cada tipo: anti-*C4A*, anti-*CP* y anti-*PON1*.

10 En otra realización preferida, el kit de este aspecto de la invención comprende anticuerpos secundarios o controles positivos y/o negativos. El kit además puede incluir, sin ningún tipo de limitación, tampones, soluciones de extracción de proteínas, agentes para prevenir la contaminación, inhibidores de la degradación de las proteínas, etc.

Por otro lado el kit de este aspecto de la invención puede incluir todos los soportes y recipientes necesarios para su puesta en marcha y optimización. Preferiblemente, el kit comprende además las instrucciones para llevar a cabo los métodos de la invención.

15 En una realización preferida, el kit de este aspecto de la invención comprende la asignación de un valor de diagnóstico y clasificación de los individuos, de manera que: si los individuos presentan un $VD > 3,87$ ng/mg se clasifican en el grupo de individuos con hepatocarcinoma.

20 Un **aspecto H** de la invención se refiere a un kit o dispositivo que comprende los elementos necesarios para cuantificar la cantidad de los productos expresión de los genes que se seleccionan de entre *C4A*, *FGA*, *CP* y *PON1*, o cualquiera de sus combinaciones.

Aún más preferiblemente, el kit de este aspecto de la invención comprende anticuerpos o fragmentos de los mismos específicos frente a cualquiera de las secuencias SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 9 o SEQ ID NO: 11 o frente a secuencias aminoacídicas que presenten un grado de identidad con dichas secuencias aminoacídicas de al
25 menos del 85%, típicamente de, al menos del 90%, preferiblemente de, al menos del 95%, más preferiblemente de, al menos del 98%, aún más preferiblemente de, al menos del 99%. En otra realización preferida, el kit de este aspecto de la invención comprende anticuerpos secundarios o controles positivos y/o negativos. El kit además puede incluir, sin ningún tipo de limitación, tampones, soluciones de extracción de proteínas, agentes para prevenir la contaminación,
30 inhibidores de la degradación de las proteínas, etc.

Por otro lado el kit puede incluir todos los soportes y recipientes necesarios para su puesta en marcha y optimización. Preferiblemente, el kit comprende además las instrucciones para llevar a cabo los métodos de la invención.

En otra realización preferida, el kit de este aspecto de la invención es adecuado para amplificar secuencias nucleótidas, que comprende sondas y/o cebadores diseñados a partir de las secuencias SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10 y/o SEQ ID NO: 12 así como opcionalmente todos aquellos elementos necesarios para llevar a cabo un procedimiento PCR.

En una realización preferida, el kit de este aspecto de la invención comprende la asignación de un valor de diagnóstico y clasificación de los individuos, de manera que: si los individuos presentan un $VD > 95$, preferiblemente un $VD > 105$, más preferiblemente un $VD > 116$, y aún más preferiblemente un $VD > 120$ se clasifican en el grupo de individuos con hepatocarcinoma.

Un **aspecto I** de la invención se refiere a un soporte sólido, o chip de proteínas, que comprende al menos uno de los anticuerpos o fragmentos de los mismos específicos frente a cualquiera de las secuencias SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 9 o SEQ ID NO: 11 o frente a secuencias aminoacídicas que presenten un grado de identidad con dichas secuencias aminoacídicas de al menos del 85%, típicamente de, al menos del 90%, preferiblemente de, al menos del 95%, más preferiblemente de, al menos del 98%, aún más preferiblemente de, al menos del 99%, o cualquiera de sus combinaciones o cualquiera de sus combinaciones, para llevar a cabo el cualquiera de los métodos de la invención.

Un **aspecto J** de la invención se refiere a un soporte sólido, o chip de ADN, que comprende oligonucleótidos o microarreglos de canal único diseñados a partir de una secuencia conocida o un ARNm de al menos uno de los genes C4A, FGA, CP y PON1. Preferiblemente, comprende los oligonucleótidos capaces de detectar el ARNm de todos los genes C4A, FGA, CP y PON1.

Así, por ejemplo, las secuencias de oligonucleótidos son construidas en la superficie del chip mediante el elongamiento secuencial de una cadena en crecimiento con un sólo nucleótido utilizando fotolitografía. Así, los oligonucleótidos son anclados por el extremo 3' mediante un método de activación selectiva de nucleótidos, protegidos por un reactivo fotolábil, mediante la incidencia selectiva de luz a través de una fotomáscara. La fotomáscara puede ser física o virtual.

Así, las sondas oligonucleótidos pueden ser de entre 10 y 100 nucleótidos, más preferiblemente, de entre 20 y 70 nucleótidos, y aún más preferiblemente, de entre 24 y 30 nucleótidos. Para la cuantificación de la expresión génica, preferiblemente se emplean aproximadamente unos 40 oligonucleótidos por gen.

Preferiblemente, los oligonucleótidos presentan modificaciones en alguno de sus nucleótidos, como por ejemplo, pero sin limitarnos, nucleótidos que tengan alguno de sus átomos con un isótopo radiactivo, normalmente ^{32}P o tritio, nucleótidos marcados inmunológicamente, como por

ejemplo con una molécula de digoxigenina, y/o inmovilizadas en una membrana. Varias posibilidades son conocidas en el estado de la técnica.

La síntesis *in situ* sobre un soporte sólido (por ejemplo, vidrio), podría hacerse mediante tecnología chorro de tinta (ink-jet), lo que requiere sondas más largas. Los soportes podrían ser, pero sin limitarse, filtros o membranas de NC o nylon (cargadas), silicio, o portas de vidrio para microscopios cubiertos con aminosilanos, polilisina, aldehídos o epoxy. La sonda es cada una de las muestras del chip. El target es la muestra a analizar: ARN mensajero, ARN total, un fragmento de PCR, etc.

En otra realización preferida de este aspecto de la invención comprende oligonucleótidos o microarreglos de canal único diseñados a partir de al menos una secuencia conocida o un ARNm comprendido cualquiera de las sondas de la invención.

Un **aspecto K** de la invención se refiere al uso de cualquiera de los kits o dispositivos de la presente invención, en la obtención de datos útiles para el diagnóstico del hepatocarcinoma en un individuo con hepatitis C, preferiblemente alcohólico.

Un **aspecto L** de la invención se refiere a un medio de almacenamiento legible por un ordenador que comprende instrucciones de programa capaces de hacer que un ordenador lleve a cabo los pasos de cualquiera de los métodos de la invención.

Otro aspecto de la invención se refiere a una señal transmisible que comprende instrucciones de programa capaces de hacer que un ordenador lleve a cabo los pasos de cualquiera de los métodos de la invención.

Los términos “polinucleótido” y “ácido nucleico” se usan aquí de manera intercambiable, refiriéndose a formas poliméricas de nucleótidos de cualquier longitud, tanto ribonucleótidos (ARN) como desoxirribonucleótidos (ADN).

Los términos “secuencia aminoacídica”, “péptido”, “oligopéptido”, “polipéptido” y “proteína” se usan aquí de manera intercambiable, y se refieren a una forma polimérica de aminoácidos de cualquier longitud, que pueden ser codificantes o no codificantes, química o bioquímicamente modificados.

Una secuencia de ácido nucleico o polinucleótido puede comprender las cinco bases que aparecen biológicamente (adenina, guanina, timina, citosina y uracilo) y/o bases distintas de las cinco que aparecen biológicamente. Estas bases pueden servir para distintos propósitos, por ejemplo, para estabilizar o desestabilizar la hibridación; para estimular o inhibir la degradación de la sonda; o como puntos de unión para restos detectables o restos de apantallamiento. Por ejemplo, un polinucleótido de la invención puede contener uno o más restos de base modificados, no estándar, derivatizados, incluyendo, pero sin limitarse a, N6-metil-adenina, N6-terc-butil-bencil-

adenina, imidazol, imidazoles sustituidos, 5-fluorouracilo, 5-bromouracilo, 5-clorouracilo, 5-yodouracilo, hipoxantina, xantina, 4-acetilcitosina, 5-(carboxihidroximetil) uracilo, 5-carboximetilaminometil-2-tiouridina, 5-carboximetilaminometiluracilo, dihidrouracilo, beta-D-galactosilqueosina, inosina, N6-isopenteniladenina, 1-metilguanina, 1-metilinosina, 2,2-dimetilguanina, 2-metiladenina, 2-metilguanina, 3-metilcitosina, 5-metilcitosina, N6-metiladenina, 7-metilguanina, 5-metilaminometiluracilo, 5-metoxiaminometil-2-tiouracilo, beta-D-manosilqueosina, 5'-metoxicarboximetiluracilo, 5-metoxiuracilo, 2-metiltio-N6-isopenteniladenina, ácido uracil-5-oxiacético, wybutoxosina, pseudouracilo, queosina, 2-tiocitosina, 5-metil-2-tiouracilo, 2-tiouracilo, 2-tiouracilo, 4-tiouracilo, 5-metiluracilo (es decir, timina), éster metílico del ácido uracil-5-oxiacético, 3-(3-amino-3-N-2-carboxipropil) uracilo, (acp3)_w, 2,6-diaminopurina, y 5-propinil pirimidina. Otros ejemplos de restos de bases modificados, no estándar, o derivatizados pueden encontrarse en las Patentes de EEUU Nos. 6.001.611; 5.955.589; 5.844.106; 5.789.562; 5.750.343; 5.728.525; y 5.679.785. Además, una secuencia de ácido nucleico o polinucleótido puede comprender uno o más restos de azúcares modificados incluyendo, pero sin limitarse a, arabinosa, 2-fluoroarabinosa, xilulosa, y una hexosa.

El término "identidad", tal y como se utiliza en esta memoria, hace referencia a la proporción de nucleótidos o aminoácidos idénticos entre dos secuencias nucleotídicas o aminoácidas que se comparan. Los métodos de comparación de secuencias son conocidos en el estado de la técnica, e incluyen, aunque sin limitarse a ellos, el programa GAG, incluyendo GAP (Devereux et al., Nucleic Acids Research 12: 287 (1984) Genetics Computer Group University of Wisconsin, Madison, (WI); BLAST, BLASTP o BLASTN, y FASTA (Altschul et al., 1999. J. Mol. Biol. 215: 403-410).

En la presente invención se entiende por variante o fragmento biológicamente activo, aquellas variantes o fragmentos de los péptidos indicados que tienen un efecto fisiológico, metabólico o inmunológico igual, o presentan la misma utilidad que los descritos. Esto es, son funcionalmente equivalentes. Dichos efectos se pueden determinar mediante métodos convencionales.

A lo largo de la descripción y las reivindicaciones la palabra "comprende" y sus variantes no pretenden excluir otras características técnicas, aditivos, componentes o pasos. Para los expertos en la materia, otros objetos, ventajas y características de la invención se desprenderán en parte de la descripción y en parte de la práctica de la invención.

EJEMPLOS DE LA INVENCION

Ejemplo 1. Análisis de la utilidad de las proteínas identificadas CNP, CP, C4A, y PON1 como biomarcadores candidatos para el desarrollo de HCC en pacientes.

Material y métodos

Todos los pacientes incluidos en el estudio se trataron en el Hospital Universitario Reina Sofía de Córdoba, España. Cada paciente dio su consentimiento y el estudio se aprobó por el comité de ética local. Todos los pacientes mostraron fibrosis F4, determinada por biopsias del hígado o por elastometría transitoria (Fibroscan). El diagnóstico de HCC se confirmó por técnicas de imagen (tomografía axial computarizada o resonancia magnética) en base a criterios no invasivos (guías de práctica clínica EASL-EORTC. 2012. *J.Hepatol.* 56(4), 908-943).

Las muestras de sangre se recogieron antes de la cirugía en tubos estériles con anticoagulante K3-EDTA (9 ml K3E VACUETTE®, Greiner Bio-One, Austria), y se separó el plasma por centrifugación a 2900 g durante 5 min a 4 °C. Las alícuotas de plasma (500 µl) se tomaron y almacenaron a - 80 °C hasta su análisis.

En todos los 52 pacientes con ingesta abusiva de alcohol (definida como el consumo regular diario de etanol superior a 40 g) y VHC se incluyeron: 25 casos con HCC (grupo con tumor) 27 controles sin HCC (grupo control). Se seleccionaron al azar 3 pacientes de cada grupo para el análisis DIGE y los pacientes restantes sirvieron como cohorte de validación (mediante Western Blot y ELISA).

Las características clínicas se resumen en las tablas 1 y 2.

Tabla 1. Información clínica de los pacientes incluidos en el análisis DIGE

Grupo	Número de paciente	Edad	Sexo	AFP	Clase child-pugh	Tamaño del tumor (cm)	Tratamiento
Tumor	1	48	Masculino	9	B	4	TACE
	2	52	Masculino	34,9	A	2,8/1,2	TACE
	3	45	Masculino	6,99	B	4,5	TACE
Sin tumor	4	49	Masculino	3	C		
	5	41	Masculino	1,58	A		
	6	44	Masculino	8,7	C		

AFP, alfa-fetoproteína; TACE, quimioembolización transarterial

Tabla 2. Información clínica de los pacientes incluidos en el análisis DIGE

Grupos de pacientes		
	Tumor (n=22)	Sin tumor (n=24)
Media de edad	62,5 ± 11,7	50,0 ± 12,7
Género (masculino/femenino)	21/1	24/0

Clase Child-Pugh			
	A	9	4
	B	5	6
	C	4	10
	n/a	4	4
AFP (ng/ml)		205,9 ± 416,1	4,5 ± 2,7
	n/a	8	8

Eliminación de proteínas de alta abundancia de muestras de plasma

Las muestras de plasma se procesaron utilizando el kit de enriquecimiento proteico ProteoMiner™ (BioRad, Hercules, California, EE.UU.), que elimina las proteínas de alta abundancia y hace posible la detección de proteínas de mediana y bajo de abundancia. Este método utiliza una biblioteca combinada de hexapéptidos unidos a perlas y construcciones de proteínas hasta alcanzar la saturación. Las muestras se procesaron de acuerdo a las recomendaciones del fabricante, utilizando una cantidad inicial de proteína de 10 mg en un volumen de 200 ml. Posteriormente, las muestras se limpiaron antes del marcaje utilizando un kit de 2-D Clean-Up (GE Healthcare, Piscataway, EE.UU.). Los pellets de proteínas se resuspendieron en tampón de lisis enfriado con hielo (30 mM de Tris, 2 M tiourea, 7 M urea, 4% w/v de CHAPS) y se ajustó el pH entre 8.5-9 con 0,5 M de NaOH. La cuantificación de proteínas se realizó utilizando el kit de 2-D Quant (GE Healthcare).

Etiquetado DIGE

Las muestras se marcaron con CyDye DIGE Fluors siguiendo las recomendaciones del fabricante (GE Healthcare). Muestras de 50 µg de proteína de plasma preparado anteriormente se marcó con 400 pmol de cualquiera de Cy3 o Cy5 para la comparación en el mismo gel 2-D. Las muestras de grupos de control y tumorales se marcaron alternativamente con Cy3 y Cy5. Las reacciones de etiquetado se realizaron en hielo y oscuridad durante 30 min y después se inactivó con un exceso molar de 50 veces lisina libre durante 10 min en hielo. También se preparó un conjunto de muestras, que se marcaron con Cy2 para ser utilizadas como estándar en todos los geles de coincidencia de imagen y en el análisis estadístico transversal. Las reacciones de etiquetado Cy3 y Cy5 (50 µg de cada uno) de cada lisado y una cantidad igual (50 µg) de Cy2 estándar, se mezclaron y se compusieron en el mismo gel 2-D.

Separación de proteínas por 2-D y gel de imágenes

Tiras de 24 cm de un gradiente de pH inmovilizado (IPG; pH 3-10) se rehidrataron durante la noche en 450 µl de tampón de rehidratación (2 M Tiourea, 7 M Urea, 2% CHAPS, 0,2% DTT, Y 1% de IPG Buffer 3-10 NL-GE Healthcare). El enfoque isoeléctrico se realizó usando un dispositivo IPGphor (GE Healthcare) y siguiendo los siguientes pasos : 2 h a 150 V , 3 h a 300 V , 3 h a 600 V (gradiente), 3 h a 1000 V (gradiente) , 3 h a 8000 V (gradiente) , y 4 horas a 8000 V , las tiras se equilibraron durante 15 min en Tris 50 mM , urea 6 M , 2 % de SDS , 30 % de glicerol , que contiene 1 % de DTT , y luego durante 15 min en la mismo tampón que contiene 2,5 % yodoacetamida .

Las tiras IPD equilibradas se transfirieron a geles de poliacrilamida 12% de 18x20 cm entre placas de vidrio de baja fluorescencia, y después se cubrieron con 0,5% de agarosa con bajo punto de fusión en tampón que contiene azul de bromofenol. Los geles se procesaron utilizando el dispositivo Ettan Dalt 6 (GE Healthcare) a 4W/gel durante 45 min, y posteriormente a 15W/gel y 10 °C hasta que el frente de colorante se había escapado de la parte inferior de los geles. Los canales de cada gel Cy2, Cy3 y Cy5 se fotografiaron con un Typhoon 9400 Variable Mode Imager (GE Healthcare). Las estadísticas y la cuantificación de la expresión de proteínas para la selección de punto se realizaron con el software DeCyder (GE Healthcare).

Digestión de manchas y análisis MS

La digestión de manchas y el análisis MALDI-TOF se realizaron en las instalaciones de proteómica UCO-SCAI (Córdoba, España) miembro de la Plataforma en red de proteómica Carlos III, ProteoRed-ISCI. Las manchas de proteínas seleccionadas se recogieron automáticamente a partir del gele en una estación de investigador ProPic (Digilab Genómica Solutions Ltd., Hutingdon, Reino Unido). Los fragmentos resultantes del gel con tripsina y se depositaron automáticamente en la placa de MALDI en una estación II PropPrep (Digilab). Las muestras se analizaron por MALDI-TOF/TOF MS con el fin de obtener el MS en el 4800 *Proteomics Analyzer Spectrometer* (Applied Biosystems Inc., California, U.S.A. La identificación de proteínas se realizó por combinación de los espectros de MS y su MS/MS con la base de datos NCBIInr, usando MASCOT (MatrixScience) como motor de búsqueda. Los pasos y las condiciones se resumen en Ferrín *et al.*, 2013 (*Liver Int.* Jul 22).

Análisis de Western blot

10 µg de proteínas de las muestras de plasma se separaron en geles de poliacrilamida al 12% y se transfirieron a membranas de nitrocelulosa (Pall Life Sciences, Nueva York, EE.UU.). Estas membranas se incubaron durante 1 hora a temperatura ambiente en Tris solución salina tamponada con Tween-20 (TTBS; 1M de Tris-HCl, pH 7,6, NaCl 5 M, 0,1% de Tween 20) que contenía 5% de leche desnatada. Los anticuerpos primarios utilizados fueron anti-carboxipeptidasa N de la subunidad catalítica (dilución 1:200; sc-18966), anti-paraoxonasa

1(dilución 1:200; SC-133919) y anti-ceruloplasmina (dilución 1:200; SC-365205), todos ellos de Santa Cruz Biotechnology, California, Estados Unidos. Los *b/ots* fueron incubados con anticuerpos primarios diluidos en TTBS que contenía 1% de leche desnatada, durante la noche a 4 °C. Después del lavado con TTBS, los *b/ots* se incubaron con el anticuerpo secundario de cabra IgG-HRP (dilución 1:25000, Santa Cruz Biotechnology) durante 1 hora a temperatura ambiente. La inmunodetección se realizó utilizando el ECL Advance (GE Healthcare) y el LAS-3000 Imager instrumento (Fujifilm).

Prueba ELISA

Se utilizaron las pruebas de ELISA para cuantificar la concentración de proteínas de las proteínas del PCN, CP, C4A y PON1 en las muestras de plasma de los grupos control (24 casos) y de tumores (22 casos). Los ensayos se realizaron siguiendo las instrucciones del fabricante (USCN Life Science Inc., Wuhan, China), y los datos obtenidos se utilizaron para evaluar la sensibilidad y especificidad de los biomarcadores en el diagnóstico de CHC en pacientes alcohólicos infectados por el VHC con cirrosis.

15 Análisis estadístico

Las variables se muestran en las tablas de frecuencia o expresadas como medias y desviaciones estándar, o como mediana y rango intercuartil (IQR) para aquellas con distribución asimétrica. La prueba de Chi-cuadrado se utilizó para las frecuencias, las pruebas de T de Student o ANOVA para las variables cuantitativas, y la U de Mann-Whitney o Kruskal-Wallis para distribuciones asimétricas. El valor de umbral óptimo de biomarcadores candidatos para la identificación de los pacientes con HCC se estableció por curvas de receptor característico de funcionamiento (ROC) que eligen el umbral más bajo posible con el mayor valor predictivo negativo. La regresión logística múltiple se utilizó para explorar si una combinación de múltiples biomarcadores mejoraría la exactitud predictiva. El valor de *p* menor de 0,05 se consideró estadísticamente significativo.

25 Resultados

Análisis e identificación de proteínas diferencialmente expresadas.

Después del análisis 2-D DIGE, entre las manchas de proteínas coincidentes, 19 puntos se regularon diferencialmente en el grupo de HCC de una manera estadísticamente significativa. Fuera de los 19 puntos expresados diferencialmente detectado por análisis DIGE, 7 proteínas diferentes se identificaron por MALDI-TOF/TOF MS como carboxipeptidasa-N (CPN), ceruloplasmina (CP), componente del complemento 4A (C4A), fibrinógeno alfa (FBG-a), inmunoglobulina (Ig) región de la cadena mu C (IGHM), albúmina de suero (ALB), y suero paraoxonasa/arilesterasa 1 (PON1). Los puntos expresados diferencialmente y el cambio de

expresión en HCC en comparación con las muestras de control se muestran en la figura 1 y la tabla 3.

Tabla 3. Proteínas identificadas diferencialmente expresadas en muestras de plasma de pacientes alcohólicos infectados por el VHC con HCC. * La significación estadística se determinó por *t*-test, donde los valores de *p* fueron $\leq 0,05$.

Spot Nº	Valor p	Veces (cambio)	Nombre de la proteína	Nº de acceso	Puntuación de la proteína	Péptidos contados	Función de la proteína	Clasificación
598	0,023	1,88	Región IgM	gi:4467842	187	15	De unión a antígeno	Respuesta Inmune
625	0,049	2,33		gi:7439150	545	19		
627	0,039	1,88		gi:41388180	553	19		
633	0,027	2,09		gi:41388180	560	19		
634	0,050	1,84		gi:7439150	460	17		
641	0,041	1,80		gi:7439150	538	18		
642	0,031	1,81		gi:7439150	582	19		
987	0,044	1,67	CPN	gi:146386938	155	9	Proteasa	Respuesta de defensa
1081	0,040	1,78	PON1	gi:190194	238	9	Hidrolasa	
495	0,020	1,86	FibrinógenoA	gi:237823914	385	23	Homeostasis, activación de plaquetas, matrix extracelular	Respuesta de fase aguda
899	0,049	1,43		gi:237823914	568	27		
1091	0,041	1,47		gi:237823914	265	16		
443	0,014	1,50	Ceruloplasmina	gi:1620909	494	23	Hierro/cobre homeostasis	
451	0,037	1,51		gi:221042622	660	28		
1186	0,050	1,98	C4A	gi:34782950	569	14	Componente del complemento	
1209	0,027	1,86		gi:34782950	375	12		
762	0,038	2,02	Albumina	gi:31615331	464	28	Proteína portadora	Regulación de la presión osmótica coloidal de la sangre
763	0,025	2,45		gi:168988718	205	13		
765	0,037	2,05		gi:168988718	835	29		

Confirmación de DIGE-resultados por Western-blot

Con el fin de verificar los resultados DIGE/MS, las muestras de control/tumorales utilizadas previamente se analizaron por Western-blot (figura 2). Se confirmó el cambio relativo del nivel plasmático de expresión de las proteínas identificadas CPN, CP, y PON1.

Validación de las proteínas identificadas como biomarcadores candidatos de desarrollo de HCC.

Para evaluar la utilidad de las proteínas identificadas CPN, CP, C4A, y PON1 como biomarcadores candidatos para el desarrollo de HCC en pacientes infectados por el VHC alcohólicos, se analizaron 46 muestras de plasma (24 del grupo control y 22 del grupo de tumores) por ELISA (Figura 3). A diferencia del análisis DIGE, las concentraciones plasmáticas de las proteínas identificadas CPN, CP, C4A, y PON1 no se expresaron diferencialmente entre el grupo control y con tumor. Por el contrario, la validación por ELISA confirmó la expresión diferencial de C4A en pacientes con HCC ($2,4 \pm 1$ ng/mg de proteína de plasma) comparados con pacientes sin HCC ($1,8 \pm 0,6$ ng/mg; $p = 0,029$). Aunque existieron concentraciones más altas de CP en los pacientes de HCC ($21,8 \mu\text{g/mg}$ [IQR 14-27]) que en los controles ($16,1 \mu\text{g/mg}$ [IQR 11-24]), estas diferencias no fueron marginalmente significativas ($p = 0,071$). La concentración en suero de PON-1 fue de $8,3 \mu\text{g/mg}$ [IQR 6-13] en pacientes con HCC y $11,5 \mu\text{g/mg}$ [IQR 7-13] en los controles ($p = 0,31$). La concentración sérica de PCN fue similar en pacientes con HCC en comparación con los controles sin HCC ($45,5 \pm 16,5$ ng/mg frente a $44,9 \pm 20,1$ ng/mg, $p = 0,91$). En la regresión logística multivariante C4A se comportó como el único factor independiente predictivo de HCC con un OR = 2,44 (IC del 95%: 1,01-5,09; $p = 0,049$). Sin embargo, la capacidad de predicción mejoró con la combinación de C4A, CP y PON1 a través de la siguiente fórmula (I): $(\text{C4A} \cdot \text{CP}) / \text{PON1}$.

C4A, CP y PON1 tal y como se usan en la fórmula arriba se refiere a los niveles de expresión en suero del paciente de cada uno de los productos de expresión de cada uno de dichos genes.

El área bajo la curva ROC de esta combinación, según la fórmula (I), para identificar a los pacientes con HCC fue de 0,75. Este valor fue más alto que el área bajo la curva ROC de C4A sola (figura 3). Con un umbral de 3,87 para la combinación de C4A, ceruloplasmina y PON1 tuvo una sensibilidad del 72% y una especificidad del 71%.

La concentración sérica de C4A, fue significativamente superior en los pacientes con HCC multinodular ($3,05 \pm 1,17$ ng / ml) en comparación con los pacientes con HCC uninodular ($1,75 \pm 0,74$ ng / ml) ($p = 0,004$).

Los datos estadísticos mostraron que el mejor punto de corte de la concentración de C4A para discriminar entre HCC uninodular y multinodular fue de 2,07 ng/mg de proteína plasmática, con una sensibilidad del 91%, y una especificidad del 88%. Este resultado quiere decir que un paciente con un nivel de C4A superior a 2,07 ng/mg hace altamente probable la existencia en el paciente de un HCC multinodular.

Los biomarcadores restantes, incluyendo la combinación de C4A, PC y PON1 a través de la fórmula propuesta, no estaban relacionados con el número de nódulos de HCC. Ninguno de ellos (C4a incluido) estuvo correlacionado con el tamaño los de nódulos.

La presencia de un mayor número de nódulos de HCC se relaciona con un estadio más avanzado del HCC. Por tanto, una mayor concentración de C4A en plasma de pacientes con HCC se asocia con un estadio tumoral más avanzado relacionado con el HCV y el alcohol, y por consiguiente, con un peor pronóstico y opciones de tratamiento más restringidas.

5

Ejemplo 2. Análisis de la utilidad de las proteínas identificadas CPN, CP, C4A, FGA, IgM y PON1 como biomarcadores candidatos para el desarrollo de HCC en pacientes

Materiales y métodos

A partir de una base de datos de pacientes con enfermedad hepática en fase terminal, recogida en el Hospital Reina Sofía Universidad de Córdoba (España) de forma prospectiva desde 2005-2013, se seleccionaron 52 pacientes con ingesta abusiva de alcohol (definido como consumo de etanol diario regular superior a 40 g) e infección por VHC. Entre los pacientes, 25 tenían HCC (grupo con tumor) y 27 estaban libres de HCC (grupo control). Se seleccionaron al azar 3 pacientes de cada grupo emparejados por edad, sexo y clasificación Child-pugh para el análisis DIGE y los pacientes restantes sirvieron como cohorte de validación (mediante Western Blot y ELISA). No existieron criterios de exclusión para el presente estudio. Se incluyeron pacientes con diferentes características del tumor con el fin de identificar biomarcadores de HCC independientemente de la progresión del tumor.

Todos los pacientes mostraron fibrosis F4, ya sea por biopsia del hígado o elastometría transitoria (Fibroscan). El diagnóstico de HCC se confirmó por técnicas de imagen (tomografía axial computarizada o resonancia magnética) sobre la base de criterios no invasivos (guías de práctica clínica EASL-EORTC; EASL-EORTC clinical practice guidelines (2012): management of hepatocellular carcinoma. *J. Hepatol.* 56: 908-943). Las características basales de los pacientes incluidos se muestran en la Tabla 4.

Tabla 4. Información clínica de los pacientes incluidos en los análisis DIGE y ELISA. Los datos se expresan como la media $\pm$ desviación estándar; Los valores-*p* del análisis univariante están incluidos en la corte de validación

VARIABLE	Análisis DIGE		Validación ELISA		<i>p</i>
	GRUPO HCC (n=3)	GRUPO CONTROL (n=3)	GRUPO HCC (n=22)	GRUPO CONTROL (n=24)	
Edad	48,3 $\pm$ 3,5	44,7 $\pm$ 4,0	62,5 $\pm$ 11,7	50,0 $\pm$ 12,7	0,001

Género; n (%)					
Masculino	3 (100%)	3 (100%)	21 (95,5%)	24 (100%)	0,47
Femenino	0 (0%)	0 (0%)	1 (4,5%)	0 (0%)	
Child-Pugh; n (%)					
A	1 (33,3%)	1 (33,3%)	13 (59,1%)	5 (20,8%)	0,008
B-C	2 (66,7%)	2 (66,7%)	9 (40,9%)	19 (79,2%)	
MELD	12,3 ± 3,2	15,3 ± 4,0	12,3 ± 5,8	15,5 ± 5,9	0,072
AFP* (ng/ml)	17,0 ± 15,6	4,4 ± 3,8	32,9 (IQR 5,3-258,6)	4,8 (IQR 2,1-8,4)	0,050
Números de nódulos de HCC	1,3 ± 0,6	-	1,77 ± 1	-	-
Números de nódulos de HCC					
Uninodular	2 (66,7%)	-	11 (50%)	-	-
Multinodular	1 (33,3%)	-	11 (50%)	-	-
Diámetro del nódulo principal (mm)	31,2 ± 14,7		37,7 ± 20		

Las muestras de sangre se recogieron de acuerdo con nuestro protocolo en tubos estériles con anticoagulante K3-EDTA (9 ml K3E VACUETTE®, Greiner Bio-One, Austria) y el plasma se separó mediante centrifugación a 2900 g durante 5 min a 4 °C. Las muestras de plasma se tomaron y se almacenaron inmediatamente a -80 °C hasta su análisis. Las muestras se transfirieron al Biobanco del Sistema Sanitario Público de Andalucía (BBSSPA) desde el Hospital Universitario Reina Sofía. Se obtuvo un consentimiento informado de cada paciente participante en el estudio y los procedimientos estuvieron supervisados por el comité de ética del Hospital Universitario Reina Sofía.

10 ***Eliminación de proteínas de alta abundancia de muestras de plasma***

Las muestras de plasma se procesaron utilizando el kit de enriquecimiento proteico ProteoMiner™ (BioRad, Hercules, California, EE.UU.), que elimina las proteínas de alta abundancia y hace posible la detección de proteínas de mediana y baja abundancia. Las muestras se procesaron de

acuerdo a las recomendaciones del fabricante, utilizando una cantidad inicial de proteína de 10 mg en un volumen de 200 µl. Posteriormente, las muestras se limpiaron antes del marcaje utilizando un kit de 2-D Clean-Up (GE Healthcare, Piscataway, EE.UU.). Los pellets de proteínas se resuspendieron en tampón de lisis enfriado con hielo (30 mM de Tris, 2 M tiourea, 7 M urea, 4% w/v de CHAPS) y se ajustó el pH entre 8.5-9 con 0,5 M de NaOH. La cuantificación de proteínas se realizó utilizando el kit de 2-D Quant (GE Healthcare).

Etiquetado DIGE, Separación de proteínas por 2-D y gel de imágenes

Las muestras se marcaron con CyDye DIGE Fluors (GE Healthcare). 50 µg de proteína de plasma de cada una de las muestras del grupo control y tumor se marcaron con Cy3 o Cy5 para su comparación en el mismo gel 2-D. También se preparó una mezcla del conjunto de muestras totales incluidas en el estudio, que se marcó con Cy2 para ser utilizada como estándar en todos los geles y servir como imagen de coincidencia y en el análisis estadístico transversal. Las reacciones de marcaje con Cy3 y Cy5 (50 µg de cada uno) de cada lisado y una cantidad igual (50 µg) de Cy2 estándar, se mezclaron y se cargaron en el mismo gel 2-D. Las condiciones del protocolo de isoelectroenfoque y electroforesis en 2-D se resumen en Ferrín et al., 2013. *Liver International*, doi: 10.1111/liv.12277. Los canales de cada gel Cy2, Cy3 y Cy5 se fotografiaron con un Typhoon 9400 Variable Mode Imager (GE Healthcare). La estadística y la cuantificación de la expresión de proteínas para la selección se realizó con el software DeCyder (GE Healthcare).

Digestión de proteínas y análisis MS

La digestión de proteínas y el análisis MALDI-TOF se realizaron en las instalaciones de proteómica UCO-SCAI (Córdoba, España) miembro de la Plataforma en red de proteómica Carlos III, ProteoRed-ISCI. Las manchas de proteínas seleccionadas se recogieron automáticamente a partir del gel y los fragmentos resultantes del gel fueron digeridos con tripsina y analizados por MALDI-TOF/TOF MS. La identificación de proteínas se realizó por combinación de los espectros de MS y su MS/MS con la base de datos NCBI nr, usando MASCOT (MatrixScience) como motor de búsqueda. Los pasos y las condiciones se resumen en Ferrín et al., 2013. *Liver International*, doi: 10.1111/liv.12277.

Análisis de Western blot

El análisis Western Blot sólo se realizó como un control de resultados DIGE e identificación MS. Se seleccionaron las proteínas expresadas diferencialmente carboxipeptidasa N (CPN), paraoxonasa 1 (PON1) y ceruloplasmina (CP). 10 µg de proteína de las muestras de plasma procesadas fueron separadas en un gel del 12% de SDS-PAGE, se transfirieron a membranas nitrocelulosa (Pall Life Sciences, NuevaYork,EE.UU.) y se incubaron con una dilución 1:200 de anticuerpos primarios contra CPN (anticuerpo policlonal de cabra; sc-18966), PON1 (conejo

anticuerpo policlonal; sc-133919) y CP (anticuerpo monoclonal de ratón; sc-365205) (Santa Cruz Biotechnology, California, U.S.A.). La inmunodetección se visualizó por quimioluminiscencia, usando ECL Avance (GE Healthcare) y el instrumento Imager LAS-3000 (Fujifilm).

5 **Ensayo ELISA**

Las pruebas de ELISA se utilizaron para cuantificar la concentración de las proteínas CPN, CP, C4A, FGA, PON1 (USCN Life Science Inc., Wuhan, China) e inmunoglobulina M (IgM) (Abcam, Cambrige, UK) en las muestras de plasma de los grupos control (24 casos) y de tumores (22 casos). Los ensayos se realizaron siguiendo las instrucciones del fabricante, y los datos obtenidos se utilizaron para evaluar la sensibilidad y especificidad de los biomarcadores en el diagnóstico de HCC en pacientes alcohólicos infectados por el VHC con cirrosis.

Análisis estadístico

Las variables se muestran en las tablas de frecuencia o expresadas como medias y desviaciones estándar, o como mediana y rango intercuartil (IQR) para aquellas con distribución asimétrica. La prueba de Chi-cuadrado se utilizó para las frecuencias, las pruebas de T de Student o ANOVA para las variables cuantitativas, y la U de Mann-Whitney o Kruskal-Wallis para distribuciones asimétricas. El valor de umbral óptimo de biomarcadores candidatos para la identificación de los pacientes con HCC se estableció por curvas de receptor característico de funcionamiento (ROC) que eligen el umbral más bajo posible con el mayor valor predictivo negativo. La regresión logística múltiple se utilizó para explorar si una combinación de múltiples biomarcadores mejoraría la exactitud predictiva. El valor de p menor de 0,05 se consideró estadísticamente significativo.

Resultados

De los 52 pacientes incluidos, 3 pacientes con HCC y 3 pacientes sin HCC se seleccionaron al azar para el análisis DIGE. Las características de estos pacientes se describen en la tabla 1 (izquierda). El resto de pacientes sirvieron como cohorte validación incluyendo 22 de pacientes del grupo HCC, y 24 pacientes del grupo control (tabla 1 derecha). La cohorte de validación de los pacientes con HCC fue mayor y mostró mejor función hepática que los pacientes sin HCC. Así, 13 de los 22 pacientes con HCC (59,1%), y sólo 5 de los 24 pacientes del grupo control (20,8%) fueron Child A ($p=0,008$). La puntuación MELD se redujo ligeramente en pacientes con HCC ($12,3 \pm 5,8$ vs $15,5 \pm 5,9$; $p=0,072$). La media de edad de los paciente con HCC fue de $62,5 \pm 11,7$ frente a $50 \pm 12,7$ en los controles ($p = 0,001$). Así, tanto la edad como la función hepática se consideraron como posibles factores de confusión, y se controlaron en el análisis de regresión logística multivariante como se describe

a continuación. En el grupo HCC el número medio de nódulos fue de $1,77 \pm 1$, siendo el 50% de ellos pacientes con HCC uninodular. El diámetro del nódulo principal fue de $37,7 \pm 20$ mm.

5 *Análisis e identificación de proteínas diferencialmente expresadas.* Después del análisis 2-D DIGE, entre las manchas de proteínas coincidentes, 19 de ellas se encontraban diferencialmente expresadas en el grupo de HCC, de una manera estadísticamente significativa. De los 19 puntos expresados diferencialmente en el análisis DIGE, 7 proteínas diferentes fueron identificadas mediante MALDI-TOF/TOF MS como carboxipeptidasa-N (CPN), ceruloplasmina (CP), componente del complemento 4A (C4A), fibrinógeno alfa (FGA), inmunoglobulina (Ig) región de la cadena mu C (IGHM), albúmina de suero (ALB), y suero paraoxonasa/arilesterasa 1 (PON1). Los
10 puntos expresados diferencialmente y el cambio de expresión en HCC en comparación con las muestras de control se muestran en la figura 1 y la tabla 5.

Tabla 5. Proteínas identificadas diferencialmente expresadas en muestras de plasma. La significación estadística se determinó por *t*-test, donde los valores de *p* fueron $\leq 0,05$.

Spot N°	Valor p	Veces (cambio)	Nombre de la proteína	N° de acceso	Puntuación de la proteína	Péptidos contados	Función de la proteína	Clasificación
598	0,023	1,88		gi:4467842	187	15		
625	0,049	2,33		gi:7439150	545	19		
627	0,039	1,88		gi:41388180	553	18		
633	0,027	2,09	Región IgM	gi:41388180	560	19	De unión a antígeno	Respuesta inmune
634	0,050	1,84		gi:7439150	480	17		
641	0,041	1,80		gi:7439150	538	18		
642	0,031	1,81		gi:7439150	582	19		
987	0,044	1,67	CPN	gi:146386938	155	9	Proteasa	Respuesta de
1081	0,040	1,78	PON1	gi:190194	238	9	Hidrolasa	defensa
495	0,020	1,86		gi:237823914	385	23	Homeostasis,	
889	0,049	1,43	FibrinógenoA	gi:237823914	568	27	activación de plaquetas, matrix	
1091	0,041	1,47		gi:237823914	265	16	extracelular	Respuesta de fase aguda
443	0,014	1,50	Ceruloplasmina	gi:1520909	494	23	Hierro/cobre	
451	0,037	1,51		gi:221042522	660	28	homeostasis	
1186	0,050	1,98		gi:34782950	569	14	Componente del complemento	
1209	0,027	1,86	C4A	gi:34782950	375	12		
762	0,038	2,02		gi:31615331	464	28		Regulación de la presión osmótica
763	0,025	2,45	Albumina	gi:168988718	205	13	Proteína portadora	coloidal de la sangre
765	0,037	2,05		gi:168988718	835	28		

Confirmación de DIGE-resultados por Western-blot. Con el fin de verificar los resultados DIGE/MS, se eligieron aleatoriamente 3 proteínas previamente identificadas con MS. Se confirmó el cambio relativo del nivel plasmático de expresión de las proteínas identificadas CPN, CP, y PON1 por Western-blot (figura 2).

Validación de las proteínas identificadas como biomarcadores candidatos de desarrollo de HCC. Para evaluar la utilidad de las proteínas identificadas CPN, CP, C4A, FGA, IgM y PON1 como biomarcadores candidatos para el desarrollo de HCC en pacientes alcohólicos infectados por el VHC, se analizaron 46 muestras de plasma (24 del grupo control y 22 del grupo de tumores) por ELISA (Figura 4A). El inmunoensayo confirmó la expresión diferencial de C4A en pacientes con HCC ($2,4 \pm 1$ ng/mg de proteína de plasma) comparados con pacientes sin HCC ($1,8 \pm 0,6$ ng/mg;

$p = 0,029$). Las concentraciones medias de FGA se incrementaron en los pacientes con HCC ($49,9 \pm 28$ vs 25 ± 12 mg / mg) ($p < 0,001$). Aunque existieron concentraciones más altas de CP en los pacientes de HCC ($21,8 \mu\text{g/mg}$ [IQR 14-27]) que en los controles ($16,1 \mu\text{g/mg}$ [IQR 11-24]), estas diferencias fueron no significativas marginalmente ($p = 0,071$). La concentración de PON-1 en plasma fue de $8,3 \mu\text{g/mg}$ [IQR 6-13] en pacientes con HCC y $11,5 \mu\text{g/mg}$ [IQR 7-13] en los controles ($p = 0,31$). La concentración plasmática de CPN fue similar en pacientes con HCC en comparación con los controles sin HCC ($45,5 \pm 16,5$ ng/mg frente a $44,9 \pm 20,1$ ng/mg, $p = 0,91$), como lo fue la concentración de IgM ($36,7 \pm 11$ vs $40,5 \pm 15,5$; $p = 0,34$). En la regresión logística multivariante, C4A se comportó como el único factor independiente predictivo de HCC con un (OR = 2,15 IC del 95%=1,16-4,10; $p = 0,015$), mientras que el fibrinógeno plasmático fue no significativo marginalmente ($p=0,058$) (Tabla 6).

El área bajo la curva ROC para C4A fue de 0,70. Un punto de corte de 1,77 ng de C4A/mg de proteína plasmática representó un valor apropiado para distinguir los pacientes cirróticos de pacientes con CHC, con una sensibilidad de 63% y una especificidad del 65%. Sin embargo la mejor capacidad predictiva se encontró con la combinación de C4A, FGA, CP y PON1 a través de la siguiente fórmula (II): $(\text{C4A} \cdot \text{FGA} \cdot \text{CP}) / \text{PON1}$

El área bajo la curva ROC de esta combinación, según la fórmula (II), para identificar a los pacientes con HCC fue de 0,81. Este valor fue más alto que el área bajo la curva ROC de C4A y FGA por separado (figura 4B). Con un umbral de 120 para la combinación de C4A, FGA, CP y PON1 tuvo una sensibilidad del 73% y una especificidad del 79%.

La concentración sérica de C4A, fue significativamente superior en los pacientes con HCC multinodular ($3,05 \pm 1,17$ ng / ml) en comparación con los pacientes con HCC uninodular ($1,75 \pm 0,74$ ng / ml) ($p = 0,004$). Los biomarcadores restantes, incluyendo la combinación de C4A, FGA, CP y PON1 a través de la fórmula propuesta, no estuvieron relacionados con el número de nódulos de HCC. Ninguno de ellos (C4A incluido) estuvo correlacionado con el tamaño del nódulo mayor.

Tabla 6. Análisis de regresión logística multivariante para los marcadores candidatos. El componente del complemento 4a fue el único predictor independiente de HCC después de controlar los posibles factores de confusión.

VARIABLE	OR	IC 95%	p
C4A*	2,15	1,16-4,10	0,015
VARIABLES ELIMINADAS DEL MODELO			
CPN	1	0,92-1,10	0,86
PON-1	0,98	0,77-1,23	0,85
CP	1,03	0,97-1,10	0,23
FGA	1,05	0,99-1,13	0,058
IgM	0,94	0,85-1,05	0,33
VARIABLES CONTROLADAS COMO POSIBLES FACTORES DE CONFUSIÓN			
Edad	1,1	1,02-1,17	0,01
CHILD (A vs. B-C)	0,14	0,15-1,39	0,09
MELD	1	0,85-1,18	0,95

*OR e intervalo de confianza para C4A se calculó como incremento de 0,5 ng/mg proteínas de plasma

C4A: componente del complemento 4A; CPN: carboxipeptidasa-N; PON1: Paraoxonasa/arilesterasa 1 de suero; CP: ceruloplasmina; FGA: fibrinógeno alfa; IgM región cadena mu inmunoglobulina C.

REIVINDICACIONES

1. Uso en una muestra aislada de sangre o plasma sanguíneo de un individuo con hepatitis C, de los niveles de expresión de los genes *C4A*, *FGA*, *CP* y *PON1* o de la cuantificación de los productos de expresión de los genes *C4A*, *FGA*, *CP* y *PON1*, para la obtención de datos útiles para el diagnóstico del hepatocarcinoma, donde la cuantificación del producto de expresión de dichos genes o la determinación de los niveles de expresión de dichos genes se realiza de forma simultanea o secuencial en cualquier orden y

donde por producto de expresión del gen *C4A*, se entiende la secuencia aminoácídica que presente un grado de identidad con la secuencia aminoácídica SEQ ID NO: 1 o SEQ ID NO: 2 de, al menos el 85%, o una secuencia nucleotídica que codifique para cualquiera de dichas secuencia aminoácídicas;

donde por producto de expresión del gen *FGA*, se entiende la secuencia aminoácídica que presente un grado de identidad con la secuencia aminoácídica SEQ ID NO: 5 o SEQ ID NO: 6 de, al menos el 85%, o una secuencia nucleotídica que codifique para cualquiera de dichas secuencia aminoácídicas;

donde por producto de expresión del gen *CP*, se entiende la secuencia aminoácídica que presente un grado de identidad con la secuencia aminoácídica SEQ ID NO: 9 de, al menos el 85%, o una secuencia nucleotídica que codifique para cualquiera de dichas secuencia aminoácídicas; y

donde por producto de expresión del gen *PON1*, se entiende la secuencia aminoácídica que presente un grado de identidad con la secuencia aminoácídica SEQ ID NO: 11 de, al menos el 85%, o una secuencia nucleotídica que codifique para cualquiera de dichas secuencia aminoácídicas.

2. Uso en una muestra aislada de sangre o plasma sanguíneo de un individuo con hepatitis C, de la cuantificación los productos de expresión de los genes *C4A*, *FGA*, *CP* y *PON1*, para la obtención de datos útiles para el diagnóstico del hepatocarcinoma, donde el producto de expresión de dichos genes se determina de forma simultanea o secuencial en cualquier orden y donde los productos de expresión se definen en la reivindicación 1.

3. El uso de los genes *C4A*, *FGA*, *CP* y *PON1* o de los productos de expresión de dichos genes según la reivindicación 1 o 2, donde el individuo es alcohólico.

4. Método de obtención de datos útiles para el diagnóstico del hepatocarcinoma en un individuo con hepatitis C, que comprende:

i) obtener una muestra biológica aislada de un individuo, preferiblemente sangre, más preferiblemente plasma sanguíneo; y

ii) cuantificar la cantidad de los productos de expresión de los genes: *C4A*, *FGA*, *CP* y *PON1*, donde por productos de expresión se entienden las secuencias amino-ácidas definidas en la reivindicación 1.

5. El método según la reivindicación anterior, que además comprende:

iii) hallar un valor de diagnóstico (VD) con los valores obtenidos en ii):

$$VD = (C4A \times FGA \times CP) / PON1$$

donde *C4A*, *FGA*, *CP* y *PON1* representan concentraciones de los productos de expresión de cada uno de los genes frente a la proteína total, entiendo por producto de expresión las secuencias aminoácidas definidas en la reivindicación 1; y

donde la cantidad del producto de expresión de los genes se mide en ng/mg (ng del producto de expresión del gen por mg de proteína total) para el producto de expresión del gen *C4A* y en µg/mg (µg del producto de expresión del gen por mg de proteína total) para los productos de expresión de los genes *FGA*, *CP*, *PON1*.

6. Un método de diagnóstico del hepatocarcinoma en un individuo con hepatitis C, que comprende los pasos (i)-(iii) según las reivindicaciones 4-5, y que además comprende clasificar a los individuos que presentan un $VD > 120$ en el grupo de individuos con hepatocarcinoma.

7. Un método de diagnóstico del hepatocarcinoma en un individuo con hepatitis C, que comprende los pasos (i)-(ii) según la reivindicación 4, y que además comprende: asignar los individuos del paso (i) al grupo de individuos con hepatocarcinoma uninodular cuando presentan un valor de producto de expresión del gen *C4A* $< 2,07$ ng/mg, o asignar a los individuos del paso (i) al grupo de individuos con hepatocarcinoma multinodular cuando presentan un valor de producto de expresión del gen *C4A* $> 2,07$ ng/mg;

donde el producto de expresión se define de acuerdo a la reivindicación 5.

8. El método según las reivindicaciones 4-7, donde la muestra biológica es una muestra de sangre.

9. El método según las reivindicaciones 4-8, donde la muestra biológica es una muestra de plasma, preferiblemente sanguíneo.

10. El método según las reivindicaciones 4-9, donde el individuo es alcohólico.

11. El método según cualquiera de las reivindicaciones 4-10, donde la cuantificación del producto de expresión de los genes *C4A*, *FGA*, *CP* y *PON1* se hace mediante un inmunoensayo.
12. El método según la reivindicación 11, donde el inmunoensayo es un ELISA.
- 5 13. Kit o dispositivo, que comprende los elementos necesarios para cuantificar la concentración o cantidad de los productos de expresión de los genes *C4A*, *FGA*, *CP* y *PON1*, donde el producto de expresión se define de acuerdo a la reivindicación 1, preferiblemente donde el producto de expresión se define de acuerdo a la reivindicación 5.
- 10 14. El kit o dispositivo según la reivindicación 13, donde dicho kit comprende anticuerpos o fragmentos de los mismos específicos frente a cualquiera de las secuencias SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 9 o SEQ ID NO: 11 o frente a secuencias aminoacídicas que presenten un grado de identidad con dichas secuencias aminoacídicas de al menos el 85%.
- 15 15. El kit o dispositivo según la reivindicación 13 adecuado para amplificar secuencias nucleótídicas, que comprende sondas y/o cebadores diseñados a partir de las secuencias SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10 y/o SEQ ID NO: 12 así como opcionalmente todos aquellos elementos necesarios para llevar a cabo un procedimiento PCR.
- 20 16. Uso del kit o dispositivo tal y como se define en cualquiera de las reivindicaciones 13-15, en la obtención de datos útiles para el diagnóstico del hepatocarcinoma en un individuo con hepatitis C, preferiblemente alcohólico.

FIGURAS

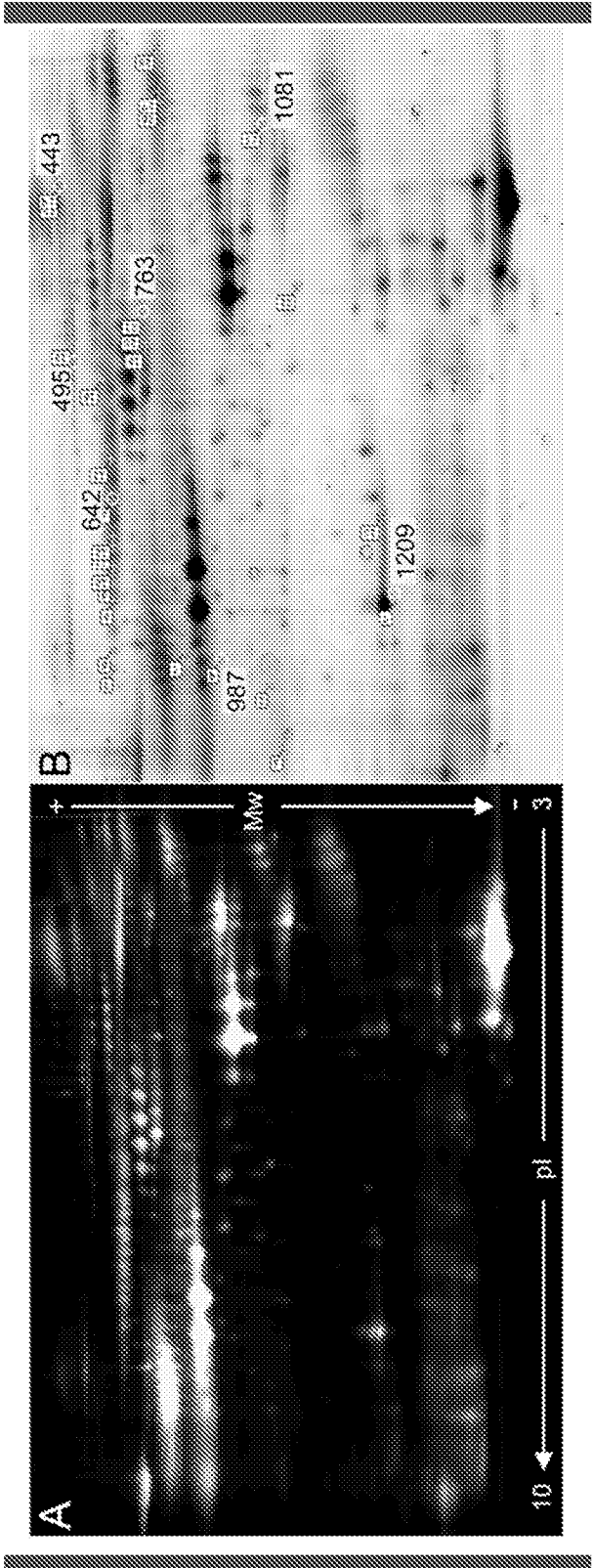


Fig.1

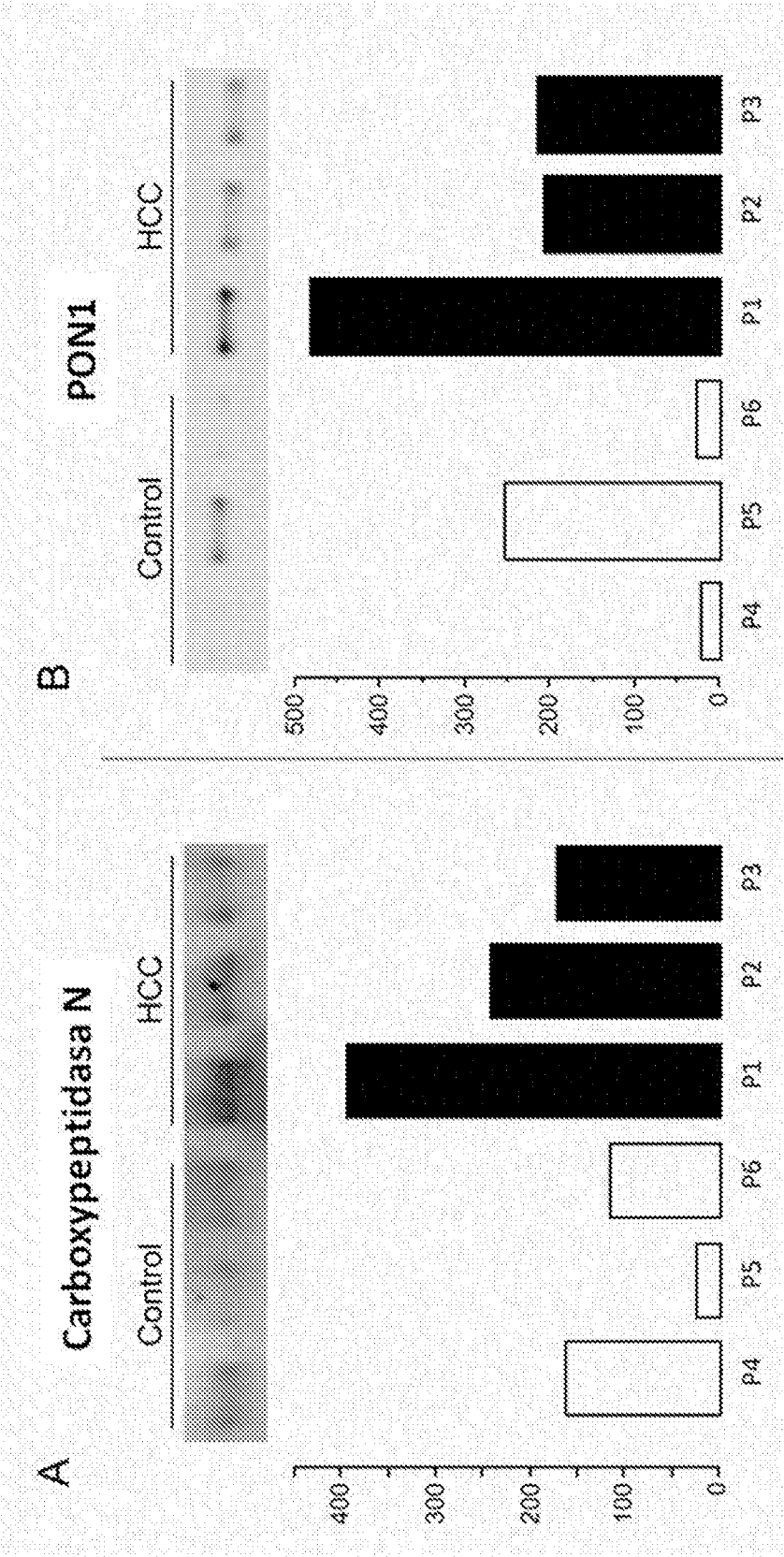


Fig.2

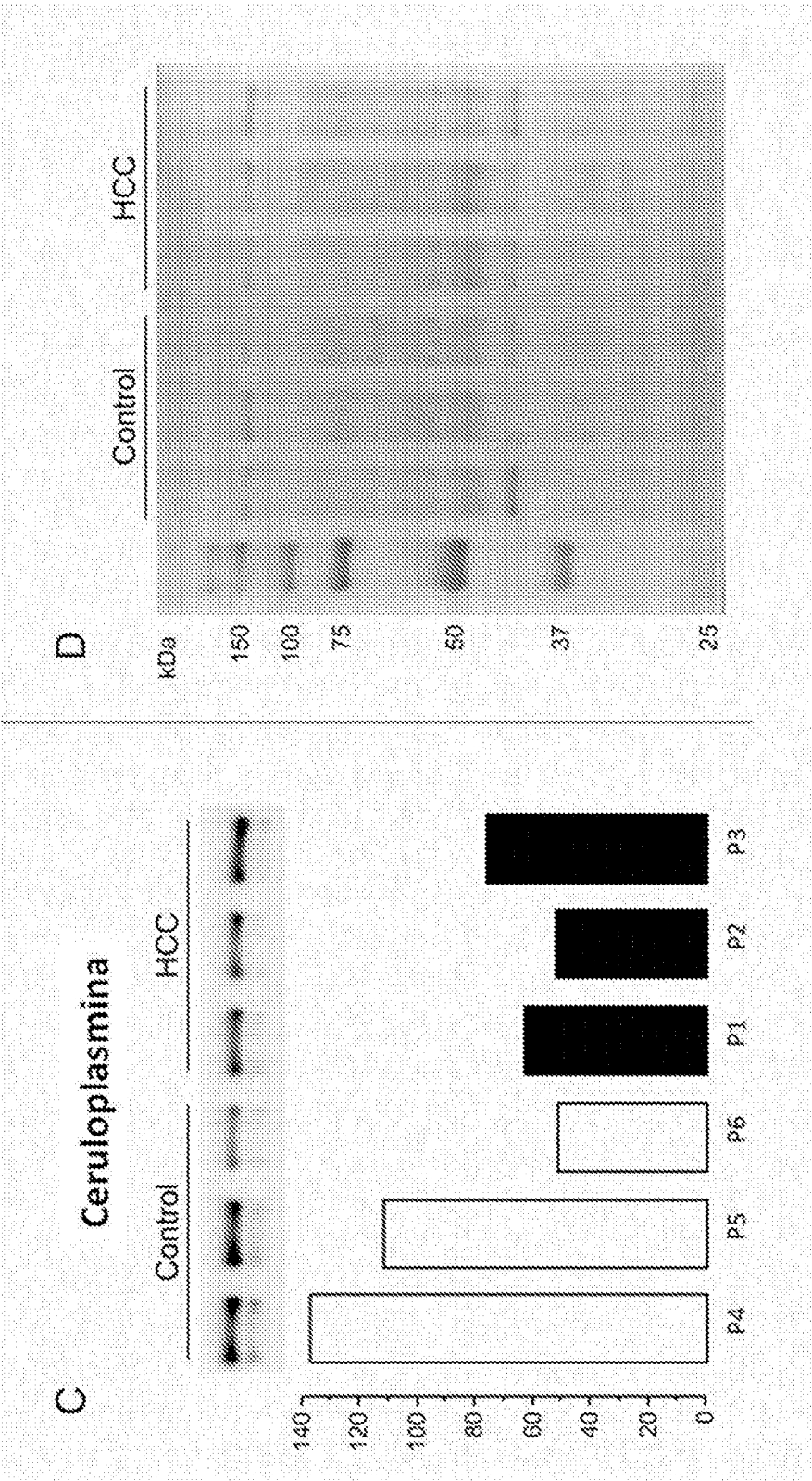
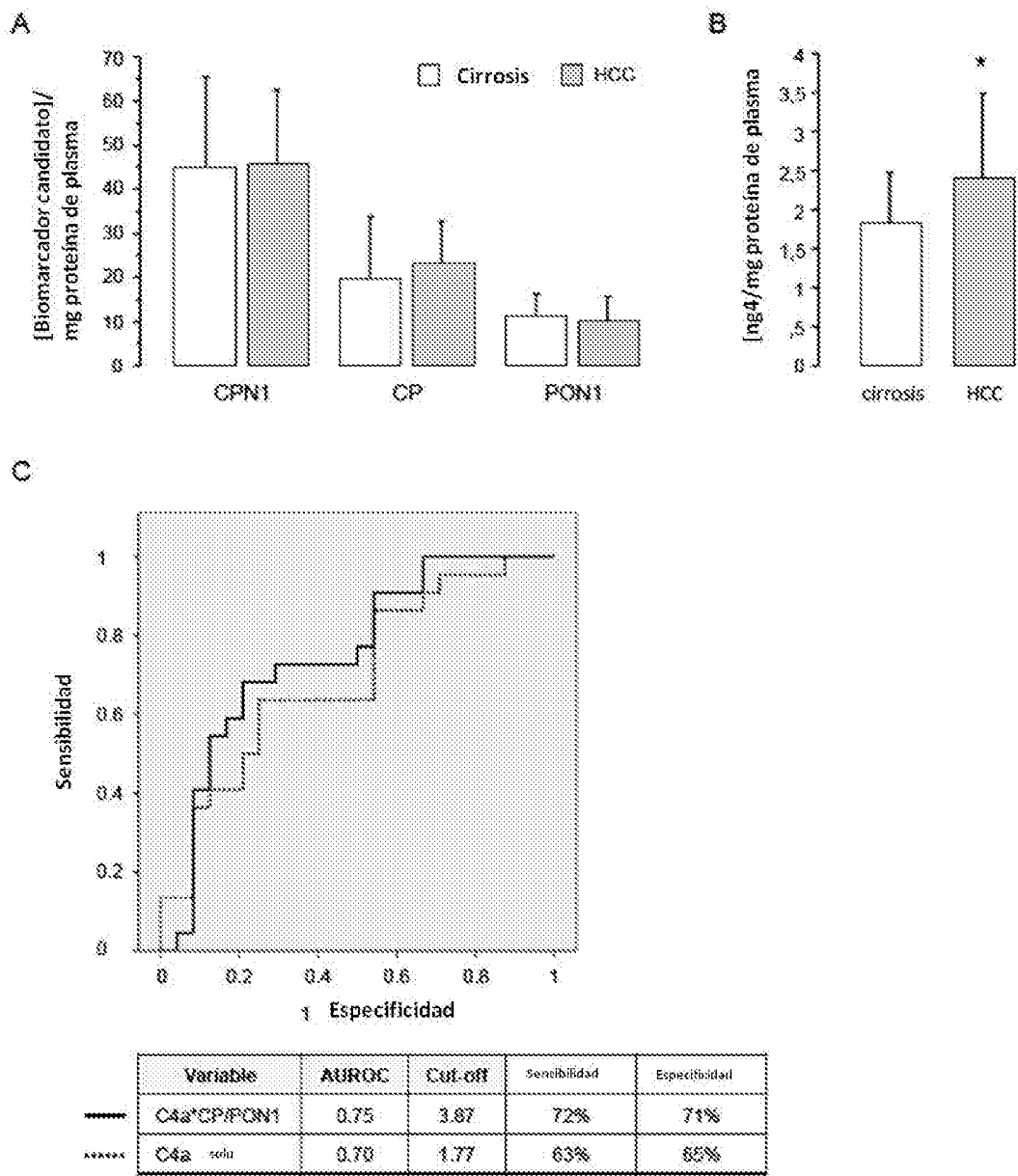


Fig.2 (cont.)

Fig.3



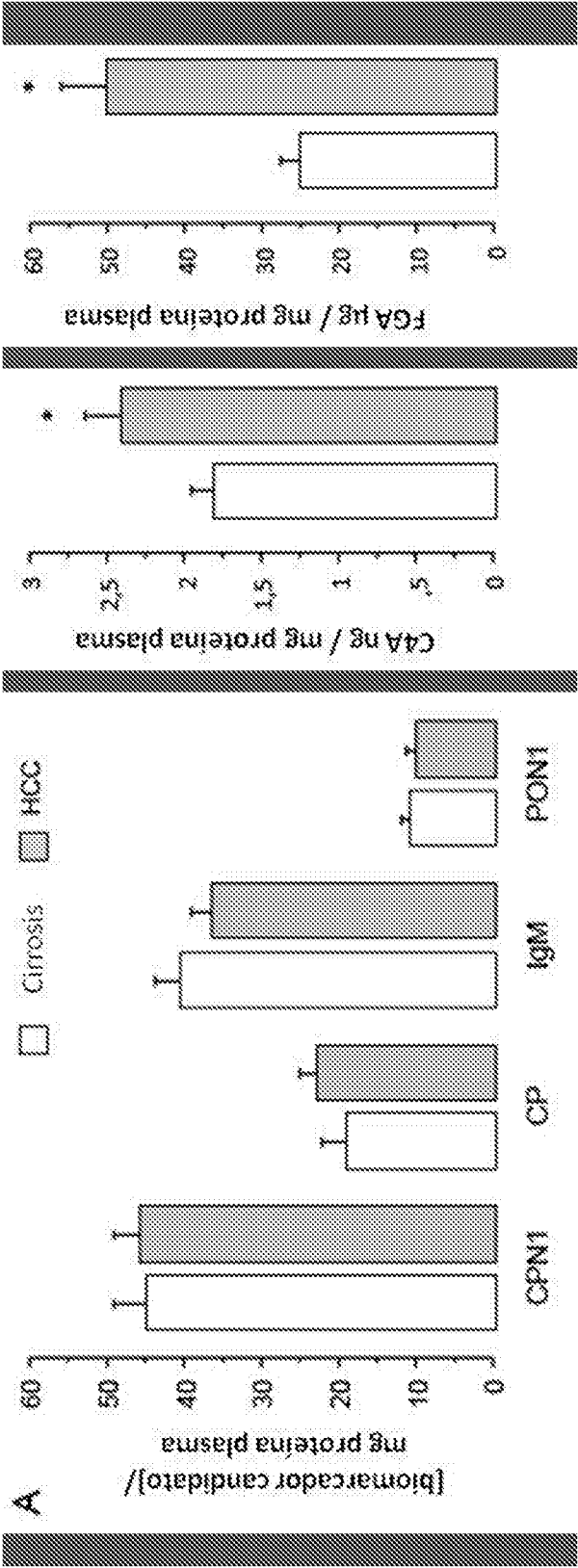
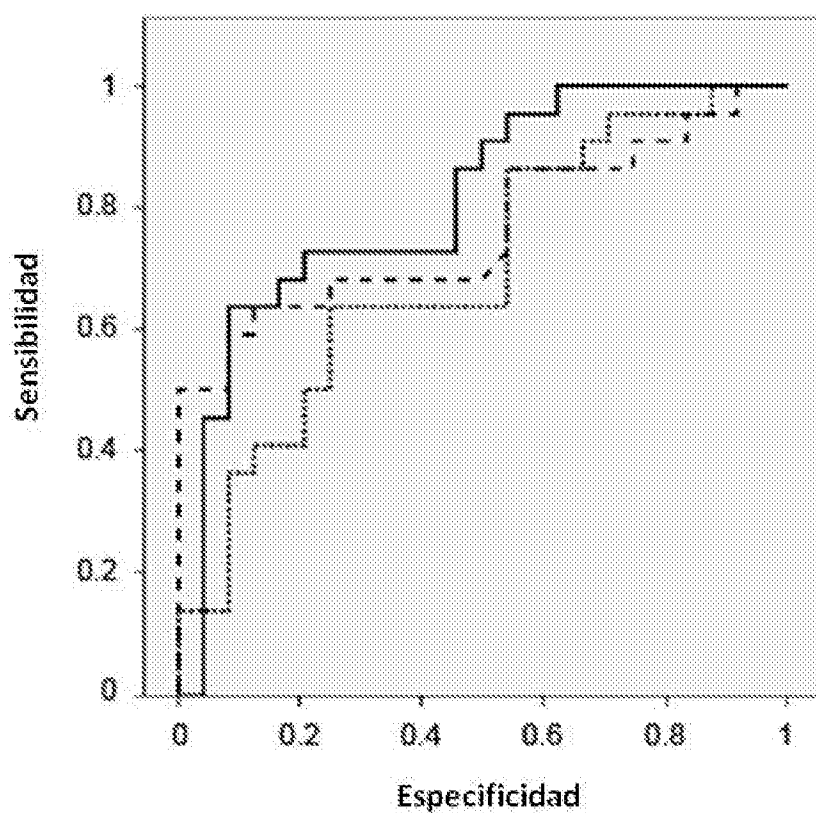


Fig. 4

B



	Variable	AUROC	Cut-off	Sensibilidad	Especificidad
————	C4A + FGA + CP/PON1	0.81	120	73%	79%
-----	C4A sólo	0.70	1.77	63%	65%
.....	FGA sólo	0.76	33.3	68%	75%

Fig. 4 (cont.)

INTERNATIONAL SEARCH REPORT

International application No.

PCT/ES2015/070402

A. CLASSIFICATION OF SUBJECT MATTER

G01N33/50 (2006.01)

G01N33/533 (2006.01)

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPODOC, INVENES, WPI, EMBASE, BIOSIS, NPL, XPESP.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	30/11/2012, KIM HYUNSOO et al. Development of biomarkers for screening hepatocellular carcinoma using global data mining and multiple reaction monitoring..PloS one United States 2013 00/00/2013 VOL: 8 No: 5 Pags: e63468 ISSN 1932-6203 (Electronic) Doi: doi:10.1371/journal.pone.0063468 pubmed:23717429	1-4, 13-16
A		5-12
X	12/06/2011, TAKESHI ISHIHARA et al. Development of quantitative plasma-glycoproteomics using label-free 2-D LC-MALDI MS and its applicability for biomarker discovery in hepatocellular carcinoma.Journal of Proteomics, 20110612 ELSEVIER, AMSTERDAM, NL 12/06/2011 VOL: 74 No: 10 Pags: 2159 - 2168 ISSN 1874-3919 Doi: doi:10.1016/j.jprot.2011.06.010 Morrice Nicholas; Burchmore Richard	1-4

☒ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance.	
"E" earlier document but published on or after the international filing date	
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"O" document referring to an oral disclosure use, exhibition, or other means.	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other documents, such combination being obvious to a person skilled in the art
"P" document published prior to the international filing date but later than the priority date claimed	"&" document member of the same patent family

Date of the actual completion of the international search
07/07/2015

Date of mailing of the international search report
(09/07/2015)

Name and mailing address of the ISA/

Authorized officer
J. Manso Tomico

OFICINA ESPAÑOLA DE PATENTES Y MARCAS
Paseo de la Castellana, 75 - 28071 Madrid (España)
Facsimile No.: 91 349 53 04

Telephone No. 91 3495583

INTERNATIONAL SEARCH REPORT

International application No.

PCT/ES2015/070402

C (continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of documents, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	31/08/2012, SUN CHUN et al. Serum paraoxonase 1 heteroplasmon, a fucosylated, and sialylated glycoprotein in distinguishing early hepatocellular carcinoma from liver cirrhosis patients. Acta Biochimica et Biophysica Sinica SEP 2012 00/09/2012 VOL: 44 No: 9 Pags: 765-773 ISSN 1672-9145(print) ISSN 1745-7270(electronic) Doi: doi:10.1093/abbs/gms055	1-4
X	JP 2006308533 A (MCBI KK ET AL.) 09/11/2006, abstract DataBase on-line. Retrieved 07.06.2015. N° Acceso 2007-061341. Retrieved from EPOQUE-WPI.	1-4

INTERNATIONAL SEARCH REPORT

International application No.

PCT/ES2015/070402

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:
- a. ☐ forming part of the international application as filed:
☒ in the form of an Annex C/ST.25 text file.
☐ on paper or in the form of an image file.
- b. ☒ furnished together with the international application under PCT Rule 13*ter*.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.
- c. ☐ furnished subsequent to the international filing date for the purposes of international search only:
☐ in the form of an Annex C/ST.25 text file (Rule 13*ter*.1(a)).
☐ on paper or in the form of an image file (Rule 13*ter*.1(b) and Administrative Instructions, Section 713).
2. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.
3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No.

PCT/ES2015/070402

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/ES2015/070402

Information on patent family members

Patent document cited in the search report	Publication date	Patent family member(s)	Publication date
JP2006308533 A	09.11.2006	JP4801930B B2	26.10.2011
-----	-----	-----	-----

INFORME DE BÚSQUEDA INTERNACIONAL

Solicitud internacional nº

PCT/ES2015/070402

A. CLASIFICACIÓN DEL OBJETO DE LA SOLICITUD

G01N33/50 (2006.01)

G01N33/533 (2006.01)

De acuerdo con la Clasificación Internacional de Patentes (CIP) o según la clasificación nacional y CIP.

B. SECTORES COMPRENDIDOS POR LA BÚSQUEDA

Documentación mínima buscada (sistema de clasificación seguido de los símbolos de clasificación)

G01N

Otra documentación consultada, además de la documentación mínima, en la medida en que tales documentos formen parte de los sectores comprendidos por la búsqueda

Bases de datos electrónicas consultadas durante la búsqueda internacional (nombre de la base de datos y, si es posible, términos de búsqueda utilizados)

EPODOC, INVENES, WPI, EMBASE, BIOSIS, NPL, XPESP.

C. DOCUMENTOS CONSIDERADOS RELEVANTES

Categoría*	Documentos citados, con indicación, si procede, de las partes relevantes	Relevante para las reivindicaciones nº
X	30/11/2012, KIM HYUNSOO et al. Development of biomarkers for screening hepatocellular carcinoma using global data mining and multiple reaction monitoring..PloS one United States 2013 00/00/2013 VOL: 8 No: 5 Pags: e63468 ISSN 1932-6203 (Electronic) Doi: doi:10.1371/journal.pone.0063468 pubmed:23717429	1-4, 13-16
A		5-12
X	12/06/2011, TAKESHI ISHIHARA et al. Development of quantitative plasma-glycoproteomics using label-free 2-D LC-MALDI MS and its applicability for biomarker discovery in hepatocellular carcinoma.Journal of Proteomics, 20110612 ELSEVIER, AMSTERDAM, NL 12/06/2011 VOL: 74 No: 10 Pags: 2159 - 2168 ISSN 1874-3919 Doi: doi:10.1016/j.jprot.2011.06.010 Morrice Nicholas; Burchmore Richard	1-4

☒ En la continuación del recuadro C se relacionan otros documentos

☒ Los documentos de familias de patentes se indican en el anexo

* Categorías especiales de documentos citados:	"T" documento ulterior publicado con posterioridad a la fecha de presentación internacional o de prioridad que no pertenece al estado de la técnica pertinente pero que se cita por permitir la comprensión del principio o teoría que constituye la base de la invención.
"A" documento que define el estado general de la técnica no considerado como particularmente relevante.	"X" documento particularmente relevante; la invención reivindicada no puede considerarse nueva o que implique una actividad inventiva por referencia al documento aisladamente considerado.
"E" solicitud de patente o patente anterior pero publicada en la fecha de presentación internacional o en fecha posterior.	"Y" documento particularmente relevante; la invención reivindicada no puede considerarse que implique una actividad inventiva cuando el documento se asocia a otro u otros documentos de la misma naturaleza, cuya combinación resulta evidente para un experto en la materia.
"L" documento que puede plantear dudas sobre una reivindicación de prioridad o que se cita para determinar la fecha de publicación de otra cita o por una razón especial (como la indicada).	"&" documento que forma parte de la misma familia de patentes.
"O" documento que se refiere a una divulgación oral, a una utilización, a una exposición o a cualquier otro medio.	
"P" documento publicado antes de la fecha de presentación internacional pero con posterioridad a la fecha de prioridad reivindicada.	

Fecha en que se ha concluido efectivamente la búsqueda internacional.
07/07/2015

Fecha de expedición del informe de búsqueda internacional.
09 de julio de 2015 (09/07/2015)

Nombre y dirección postal de la Administración encargada de la búsqueda internacional
OFICINA ESPAÑOLA DE PATENTES Y MARCAS
Paseo de la Castellana, 75 - 28071 Madrid (España)
Nº de fax: 91 349 53 04

Funcionario autorizado
J. Manso Tomico
Nº de teléfono 91 3495583

INFORME DE BÚSQUEDA INTERNACIONAL

Solicitud internacional nº

PCT/ES2015/070402

C (Continuación).		DOCUMENTOS CONSIDERADOS RELEVANTES
Categoría *	Documentos citados, con indicación, si procede, de las partes relevantes	Relevante para las reivindicaciones nº
X	31/08/2012, SUN CHUN et al. Serum paraoxonase 1 heteroplasmon, a fucosylated, and sialylated glycoprotein in distinguishing early hepatocellular carcinoma from liver cirrhosis patients. Acta Biochimica et Biophysica Sinica SEP 2012 00/09/2012 VOL: 44 No: 9 Pags: 765-773 ISSN 1672-9145(print) ISSN 1745-7270(electronic) Doi: doi:10.1093/abbs/gms055	1-4
X	JP 2006308533 A (MCBI KK ET AL.) 09/11/2006, resumen base de datos on-line. Recuperado 07.06.2015. Nº Acceso 2007-061341. Recuperado de EPOQUE-WPI.	1-4

Recuadro I Secuencia(s) de nucleótidos y/o de aminoácidos (continuación del punto 1.c de la primera hoja)

1 En lo que se refiere a **las secuencias de nucleótidos y/o de aminoácidos** divulgadas en la solicitud internacional y necesarias para la invención reivindicada, la búsqueda se ha llevado a cabo sobre la base de una lista de secuencias presentada o entregada:

a. Medios

- ☐ en papel
☒ en formato electrónico

b. Cuando

- ☐ en la solicitud internacional tal y como se presentó
☒ junto con la solicitud internacional en formato electrónico
☐ posteriormente a esta Administración a los fines de la búsqueda

2. ☐ Además, en caso de que se haya presentado más de una versión o copia de una lista de secuencias se ha entregado la declaración requerida de que la información contenida en las copias subsiguientes o adicionales es idéntica a la de la solicitud tal y como se presentó o no va más allá de lo presentado inicialmente.

3. Comentarios adicionales:

Esta Administración disponía, en el momento de la búsqueda internacional, de una lista de secuencias en formato electrónico por ella aceptada, por lo que en virtud de la Regla 13ter.1 no se requirió que el solicitante la aportara.

INFORME DE BÚSQUEDA INTERNACIONAL

Solicitud internacional n°

PCT/ES2015/070402

Recuadro II Observaciones cuando se estime que algunas reivindicaciones no pueden ser objeto de búsqueda (continuación del punto 2 de la primera hoja)

Este informe de búsqueda internacional no se ha realizado en relación a ciertas reivindicaciones según el artículo 17.2.a) por los siguientes motivos:

1. ☐ Las reivindicaciones n°s:
se refieren a un objeto con respecto al cual esta Administración no está obligada a proceder a la búsqueda, a saber:
2. ☒ Las reivindicaciones n°s: **1-4 (todas parc.)**
se refieren a elementos de la solicitud internacional que no cumplen con los requisitos establecidos, de tal modo que no pueda efectuarse una búsqueda provechosa, concretamente:
No están enteramente basadas en la descripción por lo que no se ha podido llevar a cabo una búsqueda provechosa sobre la totalidad de la materia reivindicada (art.6 del PCT). Así pues, la búsqueda se ha limitado a un método que comprende la cuantificación del conjunto de marcadores C4A, FGA, CP, PON-1, la valoración de su capacidad predictiva diagnóstica en función de la asignación de un valor de referencia, para identificar pacientes infectados por el VHC susceptibles de desarrollar hepatocarcinoma (HCC). Además la búsqueda se ha ampliado a métodos que comprendan la cuantificación de C4A y la asignación de un valor de corte, según la reivindicación 7, para discriminar entre pacientes con hepatocarcinoma uninodular o multinodular.
3. ☐ Las reivindicaciones n°s:
son reivindicaciones dependientes y no están redactadas de conformidad con los párrafos segundo y tercero de la regla 6.4(a).

Recuadro III Observaciones cuando falta unidad de invención (continuación del punto 3 de la primera hoja)

La Administración encargada de la Búsqueda Internacional ha detectado varias invenciones en la presente solicitud internacional, a saber:

1. ☐ Dado que todas las tasas adicionales requeridas han sido satisfechas por el solicitante dentro del plazo, el presente informe de búsqueda de tipo internacional comprende todas las reivindicaciones que pueden ser objeto de búsqueda.
2. ☐ Dado que todas las reivindicaciones que pueden ser objeto de búsqueda podrían serlo sin realizar un esfuerzo que justifique tasas adicionales, esta Administración no requirió el pago de tasas adicionales.
3. ☐ Dado que tan sólo una parte de las tasas adicionales requeridas ha sido satisfecha dentro del plazo por el solicitante, el presente informe de búsqueda de tipo internacional comprende solamente aquellas reivindicaciones respecto de las cuales han sido satisfechas las tasas, concretamente las reivindicaciones n°s:
4. ☐ Ninguna de las tasas adicionales requeridas ha sido satisfecha por el solicitante dentro de plazo. En consecuencia, el presente informe de búsqueda de tipo internacional se limita a la invención mencionada en primer término en las reivindicaciones, cubierta por las reivindicaciones n°s:

Indicación en cuanto a la protesta

- ☐ Se acompañó a las tasas adicionales la protesta del solicitante y, en su caso, el pago de una tasa de protesta.
- ☐ Se acompañó a las tasas adicionales la protesta del solicitante, pero la tasa de protesta aplicable no se pagó en el plazo establecido para ello.
- ☐ El pago de las tasas adicionales no ha sido acompañado de ninguna protesta.

INFORME DE BÚSQUEDA INTERNACIONAL

Informaciones relativas a los miembros de familias de patentes

Solicitud internacional nº

PCT/ES2015/070402

Documento de patente citado en el informe de búsqueda	Fecha de Publicación	Miembro(s) de la familia de patentes	Fecha de Publicación
JP2006308533 A -----	09.11.2006 -----	JP4801930B B2 -----	26.10.2011 -----